PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:		1	11) International Publication Number:	WO 00/37105
A61K 39/09, 39/40, A61P 31/04 // C07F 14/315	A2		43) International Publication Date:	29 June 2000 (29.06.00)
(21) International Application Number: PCT	US99/3(0390	(81) Designated States: AL, AM, AT, A BY, CA, CH, CN, CR, CU, CZ,	

(30) Priority Data:

60/113,048

(22) International Filing Date:

21 December 1998 (21.12.98) US

21 December 1999 (21.12.99)

- (71) Applicant: MEDIMMUNE, INC. [US/US]; 35 West Watkins Mill Road, Gaithersburg, MD 20878 (US).
- (72) Inventors: JOHNSON, Leslie, S.; 13545 Ambassador Drive, Germantown, MD 20874 (US). KOENIG, Scott; 10732 Ralston Road, Rockville, MD 20852 (US). ADAMOU, John, E.; 20822 Shamrock Glen Circle, Germantown, MD 20874 (US).
- (74) Agents: GRANT, Alan, J. et al.; Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

(57) Abstract

A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AÜ	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghaла	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary .	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	18	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		•
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
Cυ	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	Ll	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		•
EE	Estonia	LR	Liberia	SG	Singapore		
					-		
ĺ			•				

STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

5

This application is based on U.S. Provisional Application No. 60/113,048, filed 21 December 1998, which is hereby incorporated in its entirety.

10

15

FIELD OF THE INVENTION

This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species with a polypeptide comprising at least one conserved histidine triad residue (HxxHxH) and at least one helix-forming polypeptide obtained from *Streptococcus pneumoniae* as a mechanism for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic *S. pneumoniae*. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

25

In a particular aspect, the present invention relates to the prevention and treatment of pneumococcal infections such as infections of the middle ear, nasopharynx, lung and bronchial areas, blood, CSF, and the like, that are caused by pneumococcal bacteria.

BACKGROUND OF THE INVENTION

Streptococcus pneumoniae is a gram positive bacteria which is a major causative agent in invasive infections in animals and humans, such as sepsis, meningitis, otitis media and lobar pneumonia (Tuomanen et al. New Engl. J. Med. 322:1280-1284 (1995)). As part of the infective process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell et al., Micro. Path. 17:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell et al., Nature, 377:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila et al., J. Inf. Dis., 176:704-712 (1997)). A number of various other proteins have been suggested as being involved in the pathogenicity of S. pneumoniae. There remains a need for identifying polypeptides having epitopes in common from various strains of S. pneumoniae in order to utilize such polypeptides as vaccines to provide protection against a wide variety of S. pneumoniae.

25

5

10

15

20

SUMMARY OF INVENTION

In accordance with the present invention, there is provided vaccines and

vaccine compositions that include polypeptides obtained from *S. pneumoniae* and/or variants of said polypeptides and/or active fragments of such polypeptides.

The active fragments, as hereinafter defined, include a histidine triad residue(s) and/or coiled coil regions of such polypeptides.

The term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence from an alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The percent identity is determined as follows:

Percent Identity = [1- (C/R)] 100

15

20

25

10

5

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of the alignment between the Compared Sequence and the Reference Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have an aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, each being a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence in which the Percent Identity as calculated above is about

equal to or greater than a specified minimum Percent Identity than the Compared Sequence has the specified minimum Percent Identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

"Isolated" in the context of the present invention with respect to polypeptides and/or polynucleotides means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

20

5

10

15

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A-1C, respectively, report the results of three experiments using different preparations of SP36. The results demonstrate that active immunization with recombinant SP36 derived from pneumococcal strain Norway serotype 4 is able to protect mice from death in a model of pneumococcal sepsis using a heterologous strain, SJ2 (serotype 6B). In each of the three experiments shown, one hundred percent of the mice immunized

with SP36 survived for the 14-day observation period following challenge with approximately 500 cfu of pneumococci, while eighty to one hundred percent of sham-immunized mice (injected with PBS and adjuvant) died during the same period.

5

10

15

20

25

Figures 2A-2B show that passive administration of rabbit antiserum raised against Sp36 derived from Norway type 4 was able to protect mice in the pneumococcal sepsis model using two heterologous strains. Figure 2A shows that one hundred percent of the mice immunized with the SP36 antiserum survived the 21-day observation period after challenge with 172 CFU of strain SJ2 (serotype 6B). Eighty percent of the mice immunized with a control serum (rabbit anti-FimC) died by day 8, and ninety percent died by day 12. Figure 2B shows that 90 percent of the mice immunized with the Sp36 antiserum survived the 8-day observation after challenge with 862 CFU of strain EF6796 (serotype 6A). Ninety percent of the mice immunized with a control serum (collected before immunization) died by day 5.

Figure 3 is a western blot demonstrating the ability of antisera raised against recombinant Sp36 derived from strain Norway type 4 to react with Sp36 of heterologous strains. Total cell lysates were immunoblotted with mouse antisera to Sp36. A band representing Sp36 protein was detected in all 23 *S. pneumoniae* strains tested, which included isolates from each of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

Figure 4 is a Southern blot showing that the Sp36 gene from Norway type 4 hybridizes with genomic DNA from 24 other pneumococcal strains, indicating the presence of similar sequences in all these strains.

Figure 5 is a western blot showing the reactivity of patient sera with Sp36. Sp36 (either full-length, panel A; N-terminal half, panel B; or C-terminal half, panel C) was electrophoresed by SDS-PAGE and transferred to nitrocellulose. Patient sera collected soon after the onset of illness (acute serum, lanes A) or eight to 30 days later (convalescent serum, lanes C) were used to probe the blots. For patients 2, 3, and 5, convalescent serum reacted more strongly with Sp36 than did the corresponding acute serum.

10

5

Figure 6 is an amino acid alignment comparison of four related pneumococcal proteins, namely Sp36A (PhtA; SEQ ID NO:8), Sp36B (PhtB; SEQ ID NO:10), Sp36D (PhtD; SEQ ID NO:4), Sp36E (PhtE; SEQ ID NO:6), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity. Amino acid residues that match are boxed.

15

Figure 7 is a nucleotide alignment comparison of four related pneumococcal genes, namely Sp36A (PhtA; SEQ ID NO:9), Sp36B (PhtB; SEQ ID NO:11), Sp36D (PhtD; SEQ ID NO:5), Sp36E (PhtE; SEQ ID NO:7), respectively. Dashes in a sequence indicate gaps introduced to maximize the sequence similarity.

20

25

Figure 8 shows the results of immunization of mice with PhtD recombinant protein, which leads to protection from lethal sepsis. C3H/HeJ (Panel A and B) or Balb/cByJ (Panel C) mice were immunized subcutaneously with PhtD protein (15 μg in 50 μl PBS emulsified in 50 μl complete Freund's adjuvant (CFA)). The recombinant PhtD protein used in protection experiments consisted of 819 amino acid residues, starting with the cysteine

(residue 20). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 µg protein with incomplete Freund's adjuvant (IFA) was administered 3 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at week 7; and sera from each group was pooled for analysis of anti-PhtD antibody by ELISA. Mice were challenged at week 8 by an intraperitonial (i.p.) injection of approximately 550 CFU S. pneumoniae strain SJ2, serotype 6B (Panel A), 850 CFU of strain EF6796, serotype 6A (Panel B) or 450 CFU of strain EF5668, serotype 4 (Panel C). In preliminary experiments, the LD₅₀ for strain SJ2 and EF6796 were determined to be approximately 10 CFU for both strains. The LD₅₀ for strain EF5668 was determined to be < 5 CFU. Survival was determined in all groups over the course of 15 days following challenge. Data are presented as the percent survival for a total of 10 mice per experimental group. Two-sample Log-rank test was used for statistical analysis comparing recombinant Pht immunized mice to sham-immunized mice.

SUMMARY OF THE INVENTION

20

15

5

10

In accordance with one aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes at least one polypeptide that is at least 90% identical to (c) a polypeptide

sequence comprising amino acids 1-819 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 1-460 of SEQ ID NO:6 or an active fragment of the foregoing.

5

In accordance with another aspect of the present invention, there is provided a vaccine, generally in the form of a composition, that includes an active fragment of a polypeptide that is at least 90% identical to (i) a polypeptide comprising amino acids 1-800 of SEQ ID NO:8 or (ii) a polypeptide comprising amino acids 1-800 of SEQ ID NO:10.

10

The term "active fragment" means a fragment that includes one or more histidine triad residues and/or one or more coiled coil regions. A "histidine triad residue" is the portion of the polypeptide that has the sequence HxxHxH wherein H is histidine and x is an amino acid other than histidine

15

A coiled coil region is the region predicted by "Coils" algorithm: Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting Coiled Coils from Protein Sequences, *Science* **252**:1162-1164.

20

In accordance with one embodiment, the active fragment includes both one or more histidine triad residues and at least one coiled coil region of the applicable polypeptide sequence. In accordance with another embodiment, the active fragment includes at least two histidine triad residues.

25

In another embodiment, the active fragment that includes at least one histidine triad residue or at least one coiled-coil region of the applicable polypeptide includes at least about ten percent of the applicable polypeptide and no more than about 85% of the applicable polypeptide.

The polypeptide of SEQ ID NO:4 includes five histidine triad residues, as follows:

amino acids 64-69; 188-193; 296-301; 541-546; and 625-630.

5

The polypeptide of SEQ ID NO:6 includes five histidine triad residues, as follows:

amino acids 63-68; 185-190; 289-294, 376-381; and 441-446.

10

In addition, the polypeptide of SEQ ID NO:4 includes two coiled-coil regions (amino acids 120-140 and amino acids 750-772) and the polypeptide of SEQ ID NO:6 includes one coiled-coil region (amino acids 119-152).

The polypeptide of SEQ ID NO: 8 includes the following regions:

HxxHxH: amino acids 63-68, 189-194, 309-314, 550-555, 634-639. Coiled-coils: amino acids 118-145, 406-434, 462-493, 724-751.

20

In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by *S. pneumoniae*.

A vaccine, or vaccine composition, in accordance with the present invention may include one or more of the hereinabove described polypeptides or active fragments thereof. When employing more than one polypeptide or active fragment, such two or more polypeptides and/or active fragments may be used as a physical mixture or as a fusion of two or more polypeptides or active fragments. The fusion fragment or fusion polypeptide may be produced,

for example, by recombinant techniques or by the use of appropriate linkers for fusing previously prepared polypeptides or active fragments.

In an embodiment of the invention, there is provided (a) a polypeptide that is at least 95% identical or at least 97% identical or 100% identical to (i) a polypeptide sequence comprising amino acids 1 to 819 of SEQ ID NO:4 or (ii) a polypeptide sequence comprising amino acids 1-460 of SEQ ID NO:6; or (b) an active fragment of the polypeptide of (a).

In the case where the polypeptide is a variant of the polypeptide comprising the mature polypeptide of SEQ ID NO:4 or SEQ ID NO:6, or any of the active fragments of the invention, the variation in the polypeptide or fragment is generally in a portion thereof other than the histidine triad residues and the coiled-coil region, although variations in one or more of these regions may be made.

In many cases, the variation in the polypeptide or active fragment is a conservative amino acid substitution, although other substitutions are within the scope of the invention.

20

5

10

15

In accordance with the present invention, a polypeptide variant includes variants in which one or more amino acids are substituted and/or deleted and/or inserted.

25

In another aspect, the invention relates to passive immunity vaccines formulated from antibodies against a polypeptide or active fragment of a polypeptide of the present invention. Such passive immunity vaccines can be utilized to prevent and/or treat pneumococcal infections in patients. In this manner, according to a further aspect of the invention, a vaccine can be

produced from a synthetic or recombinant polypeptide of the present invention or an antibody against such polypeptide.

In still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of diseases that are caused by pneumococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or treatment of infectious diseases that are caused by *S. pneumoniae*. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of otitis media, nasopharyngeal, bronchial infections, and the like in humans by utilizing a vaccine of the present invention.

Generally, vaccines are prepared as injectables, in the form of aqueous solutions or suspensions. Vaccines in an oil base are also well known such as for inhaling. Solid forms which are dissolved or suspended prior to use may also be formulated. Pharmaceutical carriers are generally added that are compatible with the active ingredients and acceptable for pharmaceutical use. Examples of such carriers include, but are not limited to, water, saline solutions, dextrose, or glycerol. Combinations of carriers may also be used.

Vaccine compositions may further incorporate additional substances to stabilize pH, or to function as adjuvants, wetting agents. or emulsifying agents, which can serve to improve the effectiveness of the vaccine.

Vaccines are generally formulated for parental administration and are injected either subcutaneously or intramuscularly. Such vaccines can also be formulated as suppositories or for oral administration, using methods known in the art.

5

10

15

20

The amount of vaccine sufficient to confer immunity to pathogenic bacteria is determined by methods well known to those skilled in the art. This quantity will be determined based upon the characteristics of the vaccine recipient and the level of immunity required. Typically, the amount of vaccine to be administered will be determined based upon the judgment of a skilled physician. Where vaccines are administered by subcutaneous or intramuscular injection, a range of 50 to 500 µg purified protein may be given.

The present invention is also directed to a vaccine in which a polypeptide or active fragment of the present invention is delivered or administered in the form of a polynucleotide encoding the polypeptide or active fragment, whereby the polypeptide or active fragment is produced *in vivo*. The polynucleotide may be included in a suitable expression vector and combined with a pharmaceutically acceptable carrier.

In addition, the polypeptides of the present invention can be used as immunogens to stimulate the production of antibodies for use in passive immunotherapy, for use as diagnostic reagents, and for use as reagents in other processes such as affinity chromatography.

In another aspect the present invention provides polynucleotides which encode the hereinabove described polypeptides and active fragments of the invention. The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand.

In accordance with another aspect of the present invention, there is

20

25

provided

5

10

15

20

25

(A) an isolated polynucleotide that is at least 90% identical to a polynucleotide sequence encoding (i) a polypeptide comprising amino acids 1-819 of SEQ ID NO:4 or (ii) a polypeptide comprising amino acids 1-460 of SEQ ID NO:6, or

- (B) a fragment of the polynucleotide of (A) that encodes an active polypeptide fragment or
- (C) a polynucleotide that is at least 90% identical to a polynucleotide sequence encoding an active fragment of (i) a polypeptide comprising amino acids 1 to 800 of SEQ ID NO:8 or (ii) a polypeptide comprising amino acids 1 to 800 of SEQ ID NO:10.

In specific embodiments, the polynucleotide is at least 95% identical, preferably at least 97% identical, and even 100% identical to such polynucleotide sequence.

The term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of polynucleotides. The variants of the polynucleotides may be a naturally occurring allelic variant of the polynucleotides or a non-naturally occurring variant of the polynucleotides. The variants include variants in which one or more bases are substituted, deleted or inserted. Complements to such coding polynucleotides may be utilized to isolate polynucleotides encoding the same or similar polypeptides. In particular, such procedures are useful to obtain native immunogenic portions of polypeptides from different serotypes of *S. pneumoniae*, which is especially

useful in the production of "chain" polypeptide vaccines containing multiple immunogenic segments.

SEQ ID NO:5 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:4 and SEQ ID NO:7 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:6. SEQ ID NO:9 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:8, and SEQ ID NO:11 is a representative example of a polynucleotide encoding the polypeptide of SEQ ID NO:10. As a result of the known degeneracy of the genetic code, other polynucleotides that encode the polypeptides of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10 should be apparent to those skilled in the art from the teachings herein.

The polynucleotides encoding the immunogenic polypeptides described above may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptides of the present invention. The marker sequence may be, for example, a hexa-histidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptides fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

The present invention also relates to vectors which include polynucleotides encoding one or more of the polypeptides of the invention, host cells which are genetically engineered with vectors of the invention and the production of such immunogenic polypeptides by recombinant techniques in an isolated and substantially immunogenically pure form.

5

10

15

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors comprising a polynucleotide encoding a polypeptide of the invention. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the polynucleotides which encode such polypeptides. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the <u>E. coli. lac</u> or <u>trp</u>, the phage lambda P_L promoter and other promoters known to control expression of genes in

5

10

15

prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

5

10

15

20

25

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in <u>E.</u> coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the proteins.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as <u>E. coli</u>, <u>Streptomyces</u>, <u>Salmonella typhimurium</u>; fungal cells, such as yeast; insect cells such as <u>Drosophila S2</u> and <u>Spodoptera Sf9</u>; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter,

operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen, Inc.), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

10

5

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and TRP. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

20

25

15

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to

produce the gene product encoded by the recombinant sequence.

Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with

5

10

15

20

translation initiation and termination sequences. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

5

10

15

20

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

25

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

5

10

15

20

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, a french press, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art. However, preferred are host cells which secrete the polypeptide of the invention and permit recovery of the polypeptide from the culture media.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

25

The polypeptides can be recovered and/or purified from recombinant cell cultures by well-known protein recovery and purification methods. Such methodology may include ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity

chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. In this respect, chaperones may be used in such a refolding procedure. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides that are useful as immunogens in the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated.

Procedures for the isolation of the individually expressed polypeptides may be isolated by recombinant expression/isolation methods that are well-known in the art. Typical examples for such isolation may utilize an antibody to a conserved area of the protein or to a His tag or cleavable leader or tail that is expressed as part of the protein structure.

20

25

15

5

10

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a

sequence of the present invention can be obtained by direct injection of the polypeptides into an animal.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The invention will be further described with respect to the following examples; however, the scope of the invention is not limited thereby:

20 Example 1

5

10.

15

Active Protection with Anti-Sp36

A. Cloning, expression, and purification of SP36

The genomic DNA used as target for amplification was isolated from *S. pneumoniae* Norway strain (serotype 4), the same strain used for genomic sequencing. The complete sequence of the Sp36 gene (SEQ ID NO:9), and its predicted amino acid sequence (SEQ ID NO:8), are given in the Sequence Listing appended hereto. It was noted that the predicted amino acid

sequence included a hydrophobic leader sequence followed by a sequence (LSVC) similar to the consensus sequence for Type II signal peptidase (LxxC, in which both x's typically represent small amino acids). Primers (listed as SEQ ID NOS:1-3) were designed that would amplify the Sp36 gene and allow its cloning into pQE10 and expression as a histidine-tagged protein lacking the signal sequence for purification by nickel-affinity chromatography. Cloning of the fragment amplified by SEQ ID Nos 1 and 3 would result in a protein containing amino acids 2 through 800 of Sp36; cloning of the fragment amplified by SEQ ID Nos 2 and 3 would result in a protein containing amino acids 7 through 800 of Sp36 (amino acid numbers refer to SEQ ID NO:8).

B. Active Protection With Sp36 Vaccination

In each of the three experiments shown in Figures 1A-1C, C3H/HeJ mice (10/group) were immunized intraperitoneally (i.p.) with Sp36 protein (15 μg in 50 μl PBS emulsified in 50 μl complete Freund's adjuvant (CFA)). A group of 10 sham-immunized mice received PBS with adjuvant. A second immunization of 15 μg protein with incomplete Freund's adjuvant (IFA) was administered 4 weeks later; the sham group received PBS with IFA. Blood was drawn (retro-orbital bleed) at weeks 3, 6, and 9; and sera from each group were pooled for analysis of anti-Sp36 antibody by ELISA. Mice were challenged at week 10 by an i.p. injection of approximately 500 CFU *S. pneumoniae* strain SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research Hospital, Memphis, TN). In preliminary experiments, the LD₅₀ of this strain was determined to be approximately 10 CFU. Mice were monitored for 14 days for survival.

The three experiments shown in Figures 1A-1C used slightly different

5

10

15

20

preparations of recombinant Sp36. The experiments shown in Figure 1A and 1B both used Sp36 containing amino acids 20-815, but different batches of protein were used in the two experiments. The experiment shown in Figure 1C used Sp36 containing amino acids 25-815.

5

10

15

20

25

In the experiment shown in Figure 1A, 9-week sera collected from the ten mice immunized with Sp36 (first batch) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimmunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (520 cfu of pneumococci) for 14 days. Eighty percent of sham-immunized mice were dead by day 4, and the remainder survived.

In the experiment shown in Figure 1B, 9-week sera collected from the ten mice immunized with Sp36 (second batch) had an endpoint ELISA titer of >1:4,096,000. No anti-Sp36 antibody was detected in sera from shamimmunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, eighty percent were dead by day 4, and all died by day 9.

In the experiment shown in Figure 1C, 9-week sera collected from the ten mice immunized with Sp36 (containing amino acids 25- 815) had an endpoint ELISA titer of 1:4,096,000. No anti-Sp36 antibody was detected in sera from sham-immunized mice. One hundred percent of the mice immunized with Sp36 protein survived the challenge (510 cfu of pneumococci) for 14 days. Of the sham-immunized mice, ninety percent died by day 4, and all died by day 12. These data demonstrate that immunization of mice with recombinant Sp36 proteins elicits a response capable of

WO 00/37105

protecting against systemic pneumococcal infection and death. This protection was not strain-specific: the recombinant pneumococcal protein was cloned from a serotype 4 strain, while the challenge was with a heterologous strain, SJ2 (serotype 6B).

5

Example 2

Passive Protection with Anti-Sp36 Antisera

A. Generation of Rabbit Immune Sera

10

15

Following collection of preimmune serum, a New Zealand White rabbit was immunized with 250 μg of Sp36 (containing amino acids 20-815) in CFA. The rabbit was given two boosts of 125 μg Sp36 in IFA on days 29 and 50 and bled on days 39 and 60. A second rabbit was immunized with a control antigen, *E. coli* FimC.

B. Passive Protection in Mice

C3H/HeJ mice (10 mice/group) were passively immunized by two i.p. injections of 100 μl of rabbit serum. The first injection was administered twenty-four hours before challenge with 172 cfu of *S. pneumoniae* strain SJ2, and the second injection was given four hours after challenge. Figure 2 shows the survival of mice after infection with two different strains of pneumococci.

25

Figure 2A shows that of mice injected with 172 cfu of strain SJ2 (Figure 2A), one hundred percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 21-day observation period. Of the mice immunized with the control serum (anti-FimC), eighty percent

died by day 8, and ninety percent died by day 12. Figure 2B shows that of mice injected with 862 cfu of strain EF6796, ninety percent of the mice immunized with rabbit immune serum raised against Sp36 protein survived the 8-day observation period. Of those given a control serum (collected from a rabbit before immunization), ninety percent died by day 8.

These data indicate that the protection against pneumococcal infection resulting from immunization with Sp36 is antibody-mediated, since mice can be protected by passive transfer of serum from a hyperimmunized rabbit. As seen in the mouse active challenge experiments described above, serum directed against recombinant Sp36 protein cloned from a serotype 4 strain was protective against challenge with heterologous strains.

Example 3

5

10

20

25

15 Conservation of Sp36 Among Strains of S. pneumoniae

A. Western blotting

The 23 pneumococcal strains used in this experiment were obtained from the American Type Culture Collection (Rockville, MD) and include one isolate each of the 23 serotypes in the multivalent pneumococcal vaccine. For total cell lysates, pneumococci were grown to mid-logarithmic phase (optical density at 620 nm, 0.4 to 0.6) in 2 ml Todd-Hewitt broth with 0.5% yeast extract (Difco, Detroit, ME) at 37°C. Bacteria were harvested by centrifugation and washed twice with water. Pellets were resuspended in 200 µl lysis buffer (0.01% sodium dodecyl sulfate, 0.15 M sodium citrate and 0.1% sodium deoxycholate) and incubated at 37°C for 30 min, then diluted in an equal volume 2x SSC (0.3 M sodium chloride, 0.03 M sodium citrate). Lysates were separated by SDS-PAGE, transferred to nitrocellulose

membranes (Bio-Rad Laboratories, Hercules, CA), and probed with antibody in a standard Western blotting procedure. Sera from ten C3H/HeJ mice immunized with Sp36 (as described in Example 1) were pooled and used at a dilution of 1:3000. Bound antibody was detected with peroxidase-conjugated sheep anti- mouse IgG using the chemiluminescence kit from Amersham, Inc. (Cambridge, MA).

The mouse anti-Sp36 sera detected two major bands with apparent molecular weights of 97 and 100 kDa in all 23 pneumococcal lysates tested (shown in Figure 3). The Sp36 signals obtained from *S. pneumoniae* serotypes 1, 5, 17F and 22F were lower, indicating either that the level of Sp36 expression is reduced in these strains, or that Sp36 in these strains is antigenically different.

These data show that Sp36 is antigenically conserved among strains of the 23 pneumococcal serotypes represented in the current polysaccharide vaccine.

20 B. Southern blotting

5

10

Genomic DNA was prepared from each of the 23 pneumococcal strains listed in the previous section and also from strain SJ2. DNA was digested with *PvuII* and *BamHI*, electrophoresed in an agarose gel and transferred to a nylon membrane. A probe was prepared by amplifying the Sp36 gene from Norway type 4 DNA (as in Example 1) and labeling the amplified fragment with fluorescein by the random-priming method, using a kit from Amersham. Hybridization, washing, and exposure of film were carried out as in the protocol supplied by Amersham. Figure 4 shows that

the Sp36 probe hybridized with DNA from each of the 24 strains studied. The lane marked "M" contained DNA from lambda phage, digested with Hindll and labeled with fluorescein, as molecular weight markers.

5 Example 4

10

15

20

25

Immunogenicity of Sp36 in Humans

In order to determine whether Sp36 is immunogenic during human patients infection, with culture-proven pneumococcal sera from used Western bacteremia blots containing pneumococcal in were recombinant Sp36 protein. In the experiment shown in Figure 5, sera from five patients (indicated as 1 through 5) were diluted 1:3000 and used to probe blots containing full-length Sp36, the N-terminal half of Sp36 (preceding the proline-rich region), or the C-terminal half of Sp36 (following the proline-rich region). Lanes labeled A (acute) were probed with serum collected shortly after diagnosis of pneumococcal infection; lanes C (convalescent) were probed with serum collected either one month later (patients 1, 2, and 3) or eight days after the first serum collection (patients 4 and 5). For patients 2, 3 and 5, reactivity of the convalescent serum with Sp36 was stronger that that of the corresponding acute serum. difference between the acute and convalescent sera was particularly evident for reactivity with the C-terminal half of the protein.

In additional experiments (not shown), convalescent sera from 23 patients with pneumococcal infections were tested individually for reactivity with full-length Sp36: 20 of the 23 sera were found to bind Sp36 on a Western blot.

These experiments indicate that Sp36 is recognized by the human

immune system and suggest that antibodies able to bind the Sp36 protein may be produced during natural *S. pneumoniae* infection in humans. Since the patients were infected with a variety of pneumococcal strains, these data also support the idea that Sp36 is antigenically conserved.

5

Example 5

Table 1 provides the percent identity between the various sequences.

10

15

Alignment of the predicted amino acid sequences of PhtA, PhtB, PhtD, and PhtE using the MEGALIGN program of Lasergene showed strong N-terminal homology with substantial divergence of the C-termini (Figure 6). The alignment of the nucleotide sequences of the same genes is shown in Figure 7. Amino acid and nucleotide sequences were compared using the identity weighting in a Lipman-Pearson pairwise alignment, in which the number of matching residues is divided by the total of matching residues plus the number of mismatched residues plus the number of residues in gaps. In the table below, the percent identity between each pair of sequences is shown at the intersection of the corresponding row and column.

20

25

Example 6

Active Protection with PhtD Vaccination.

Mice immunized with recombinant PhtD derived from strain N4 generated potent antibody titers (reciprocal endpoint titers ranging form 2,048,00 to 4,096,000). Mice immunized with PhtD were protected against death following intraperitoneal injection with either of three heterologous strains, SJ2 (serotype 6B; provided by P. Flynn, St. Jude Children's Research

Hospital, Memphis, TN), EF6796 (serotype 6A) or EF5668 (serotype 4; both strains provided by D. Briles, University of Alabama, Birmingham). In the experiment shown in Figure 8 (Panel A), all ten of the sham-immunized mice died within 10-days after challenge with virulent pneumococci (strain SJ2), while eighty percent of the PhtD-immunized mice survived the 15-day observation period. Immunization with PhtD also protected against a serotype 6A strain, EF6796 (Panel B) and a serotype 4 strain, EF5668 (Panel C). In the experiment shown in Figure 8 (Panel B), all ten of the sham-immunized mice died within 7-days after challenge with virulent pneumococci (strain EF6796), while ninety percent of the PhtD-immunized mice survived the 15-day observation period. In the experiment shown in Figure 8 (Panel C), all ten of the sham-immunized mice died within 6-days after challenge with virulent pneumoccoci (strain EF5668), while eight of nine mice immunized with PhtD survived the 15-day observation period.

Table 1. Percent Identities

Percent Identity Between Amino Acid Sequences							
	PhtA	PhtB	PhtD	PhtE			
PhtA	~~~	66.4	63.9	49.5			
PhtB	•		87.2	49.5			
PhtD				49.8			
PhtE							
Percent Identity Between Nucleotide Sequences							
	PhtA	PhtB	PhtD	PhtE			
PhtA	***	58.3	59.3	47.9			
PhtB			86.4	47.4			
PhtD			- 	47.9			
PhtE							

WHAT IS CLAIMED IS:

5

10

15

20

1. A vaccine composition comprising:

- (a) at least one member selected from the groups consisting of (i) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-819 of SEQ ID NO:4; (ii) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-460 of SEQ ID NO:6; (iii) a fragment of the polypeptide of (i) that includes at least one of a histidine triad residue or coiled-coil region; (iv) a fragment of the polypeptide of (ii) that includes at least one of a histidine triad residue or a coiled-coil region; (v) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:8, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids; and (vi) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:10, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids; and
 - (b) a pharmaceutically acceptable carrier.
- 2. A process for preventing infection caused by *S. pneumoniae* comprising:

administering the vaccine of claim 1.

- 3. A vaccine composition comprising:
- (a) at least one antibody against a member selected from the group consisting of (i) a polypeptide comprising a polypeptide sequence that

is at least 90% identical to amino acids 1-819 of SEQ ID NO:4; (ii) a polypeptide comprising a polypeptide sequence that is at least 90% identical to amino acids 1-460 of SEQ ID NO:6; (iii) a fragment of the polypeptide of (i) that includes at least one of histidine triad residue or coiled-coil region; (iv) a fragment of the polypeptide of (ii) that includes at least one of a histidine triad residue or a coiled-coil region; (v) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:8, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids and (vi) a fragment of a polypeptide that is at least 90% identical to the polypeptide sequence comprising amino acids 1-800 of SEQ ID NO:10, wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least one of a histidine triad residue or coiled-coil region wherein said fragment includes at least 80 amino acids and no more than 680 amino acids.

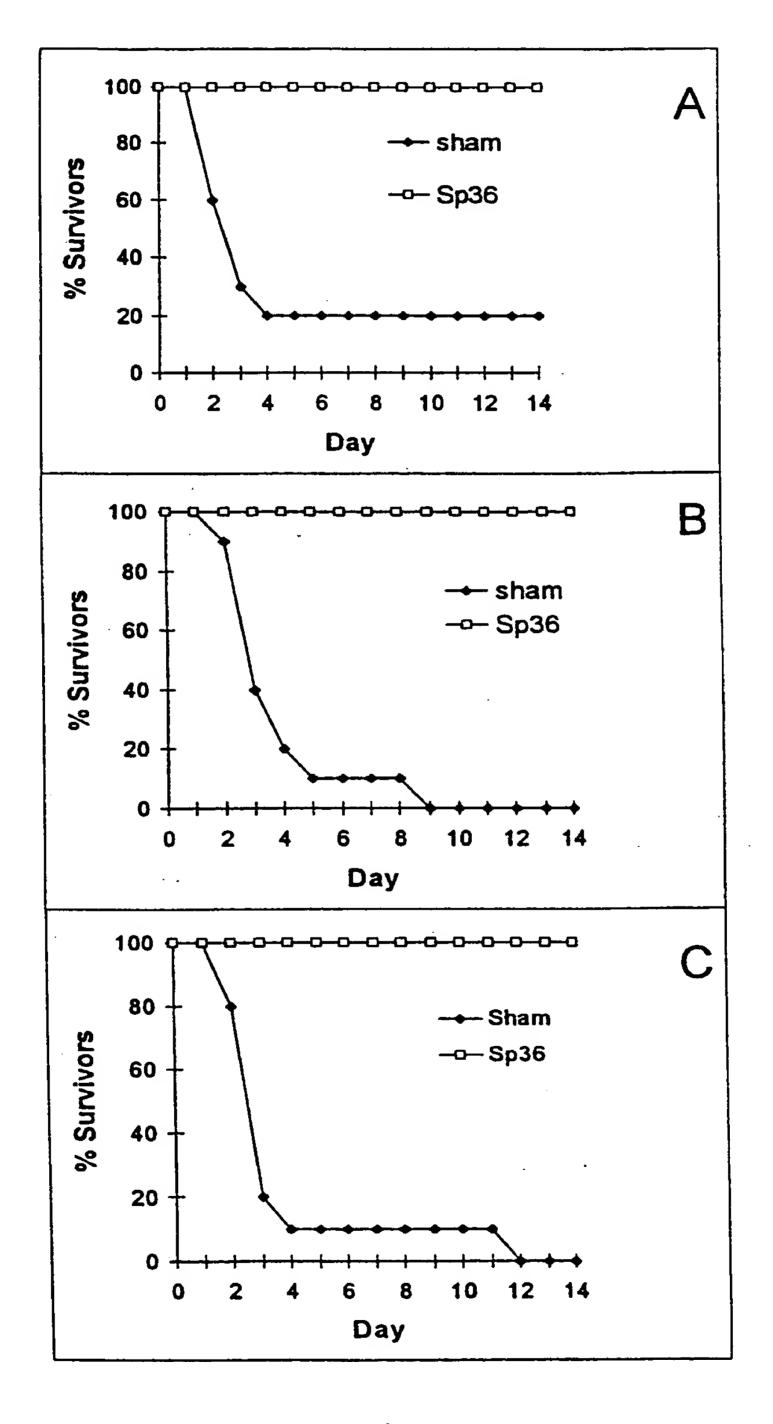
20

15

5

10.

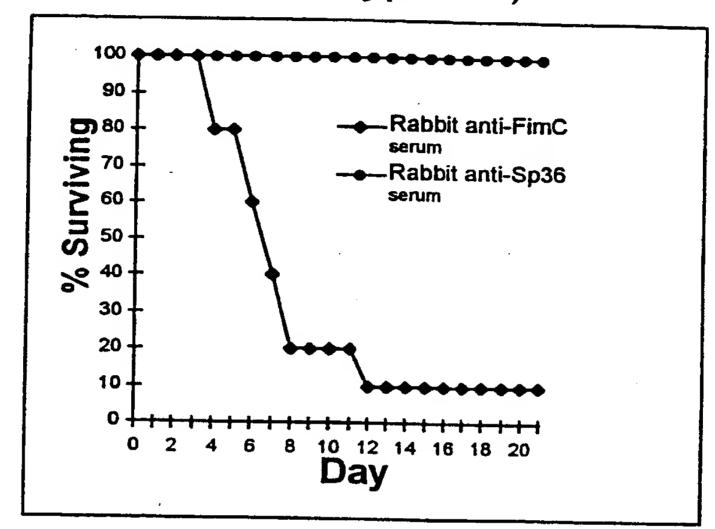
Figure 1



1/17

Figure 2

A. Strain SJ2 (serotype 6B)



B. Strain EF6796 (serotype 6A)

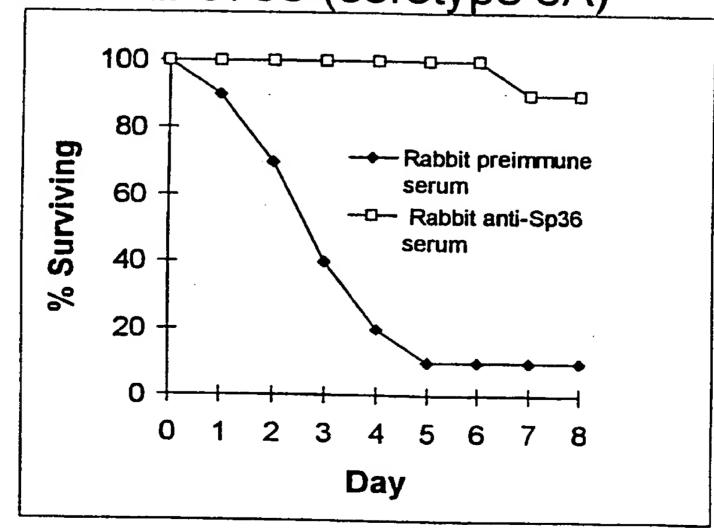


FIG.3A

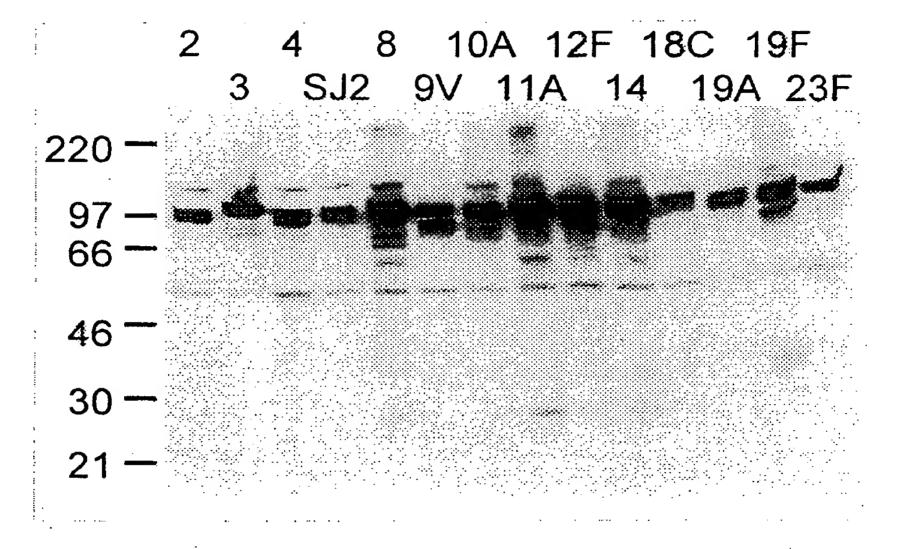
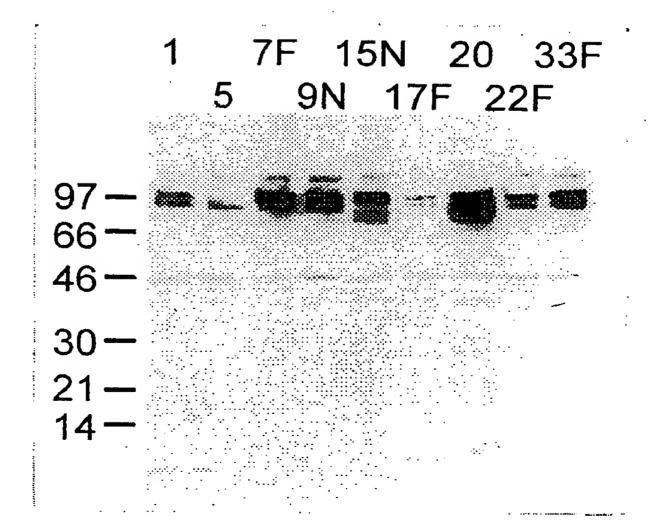


FIG.3B



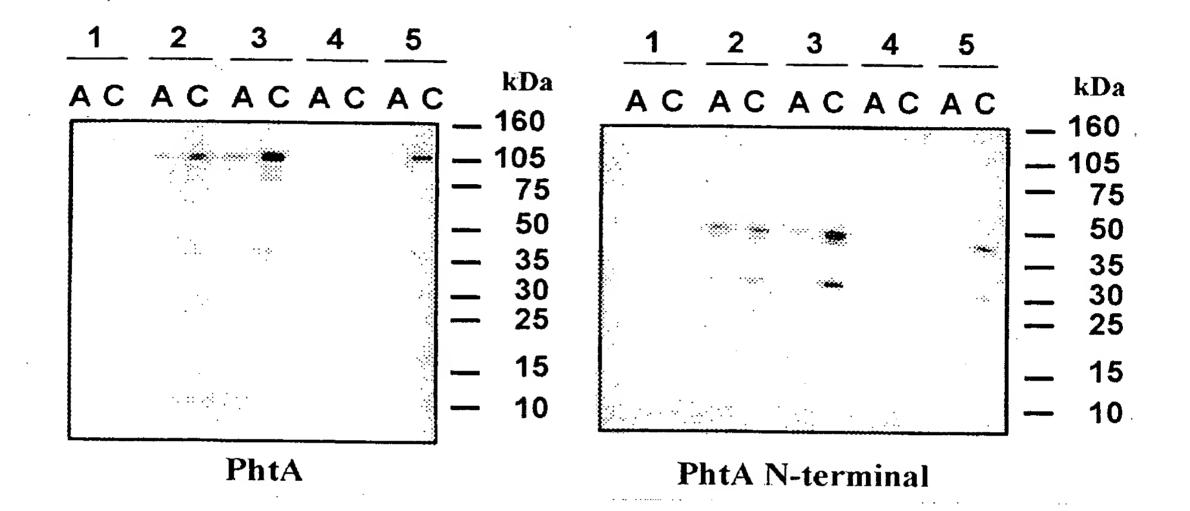
F1G.4



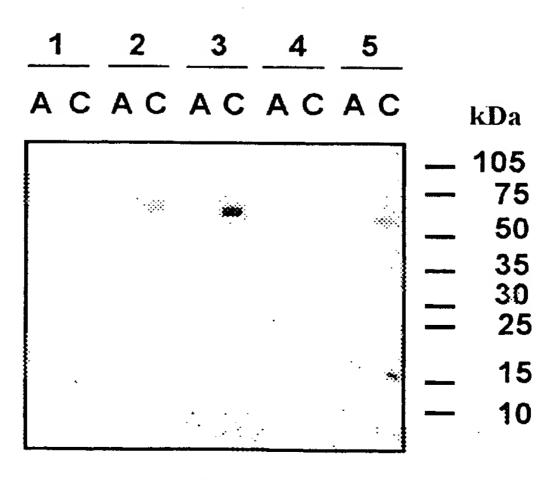
4/17
SUBSTITUTE SHEET (RULE 26)

FIG.5A

FIG. 5B



F1G.5C



PhtA C-terminal

5/17 SUBSTITUTE SHEET (RULE 26)

Figure 6(a)

CSYELGRHQAGQXKKESNRVSYIDGDQAGQKAENLTPDEVS	KREGINAEQ Hajority
10 20 30 40	50
1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVS	KREGIN-AEQ Phed. PRO
1 CSYELGRYQAGODKKESNRVAYIDGDQAGQKAENLTPDEVS	KREGINAEQ PhtB.pro
1 CSYELGLYOA-RTVKENNRVSYIDGKQATQKTENLTPDEVS	
1 CAYALHQHRS-QENKDNNRVSYVDCSQSSQKSENLTPDQVS	O KEGIQIA E O Phee. PRO
IVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLHKDP	YOLKDSDIV Majority
	100
51 IVIKITOQGYVTSHGDHYHYYNGKVPYDAIISEELLHKDP: 51 IVIKITOQGYVTSHGDHYHYYNGKVPYDAIISEELLHKDP:	YQLKDSDIV PhtB.pro
51 IVIKITOQGYVTSHGDHYHYYNGKVPYDAIISEELLHKDPR 50 IVIKITOQGYVTSHGDHYHYYNGKVPYDA <u>II</u> SEELLHKDPR	
50 IVIKITDOGYVTSHGDHYHYYNGKVPYDALPSEELLHKDP	
NEVKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINROKOE!	SANAL GG MAJORILLY
110 120 130 140	150
101 NEIKGGYVIKVDGKYYVYLKDAAHADNIRTKEEIKRQKQE	
101 NEIKGGYVIKVNGKYYVYLKDAAHADNIRTKEEIKRQKQEE	
100 NEVEGGYVIKVDGKYYVYLKDAAHADNVRTKDEINROKQEI 100 NEVEGGYIIKVDGKYYVYLKDAAHADNVRTKDEINROKQEI	
RNDXAVAAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPH	DHYHYIPKN Majority
160 170 180 190	200
149 SIN DIQIA VIVIA ARAQGRYTTD DGY I PNASDII ZDTGDAY I V PH	DHYHYIPKN Phed.PRO
148 RADNAVAAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPH	DHYHYIPKN Phrs.pro
148 RADNAVAARAQGRYTTDDGYIFNASDIIEDTGDAYIVPH	DHYHYIPKN Phts.pro
148 RADNAVAAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPH	DHYHYIPKN Phts.pro
148 RADNAVAARAQGRYTTDDGYIFNASDIIEDTGDAYIVPH	DHYHYIPKN PhtB.pro DHYHYIPKN PhtA.PRO GGHYHYIPKS Phtz.PRO
148 RADNAVAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO	DHYHYIPKN PhtB.pro DHYHYIPKN PhtA.PRO GGHYHYIPKS Phtz.PRO
148 RADNAVAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 RVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO ELSASELAAREAYLNGK	DHYHYIPKN PhtB.pro DHYHYIPKN PhtA.PRO GHYHYIPKS PhtE.PRO ANPAOPRLSE Majority 250
148 RADNAVAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN PhtB.pro DHYHYIPKN PhtA.PRO GHYHYIPKS PhtE.PRO ANPAQPRLSE Majority 250 ANPAQPRLSE PhtD.PRO ANPAQPRLSE PhtB.pro
148 RADNAVAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN Phts.pro DHYHYIPKN Phts.pro GHYHYIPKS Phts.pro ANPAQPRLSE Majority 250 ANPAQPRLSE Phtb.pro ANPAQPRLSE Phtb.pro PSVSNPGTTN Phts.pro
148 RADNAVAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN Phts.pro DHYHYIPKN Phts.pro GHYHYIPKS Phts.pro ANPAQPRLSE Majority 250 ANPAQPRLSE Phtb.pro ANPAQPRLSE Phtb.pro PSVSNPGTTN Phts.pro
148 RADNAVAAARAQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVPNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	250 ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro BYSNPGTTN Phth.pro BYSNPGTTN Phth.pro
148 RAD NA VAAARA Q GRYTTD D GY I PNASD I I EDTG DAY I V PHO 150 RND GAVALARS Q GRYTTD D GY I PNASD I I EDTG DAY I V PHO 146 RV N S NV AVARS Q GRYTTND GY V PNPAD I I EDTG NAY I V PHO ELSASELAAAEAY LNGK	DHYHYIPKN Phts.pro DHYHYIPKN Phts.pro GHYHYIPKS Phts.pro ANPAQPRLSE Majority 250 ANPAQPRLSE Phtb.pro ANPAQPRLSE Phtb.pro PSVSNPGTTN Phts.pro BTASD NN Phts.pro
148 RADNAVAAARAQGRYTTDDGYIFNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN PhtB.pro DHYHYIPKN PhtB.pro GHYHYIPKS PhtE.PRO ANPAOPRLSE Majority 250 ANPAOPRLSE PhtD.PRO ANPAOPRLSE PhtB.pro PSVSNPGTTN PhtA.PRO STASD NN PhtE.PRO LVFDPAOITS Majority 300
148 RADNA VAAARA Q G R Y T T D D G Y I P N A S D I I E D T G D A Y I V P H C 150 R N D G A V A L A R S Q G R Y T T D D G Y I F N A S D I I E D T G D A Y I V P H C 146 K V N S N V A V A R S Q G R Y T T N D G Y V P N P A D I I E D T G N A Y I V P H C E L S A S E L A A A E A Y L N G K	DHYHYIPKN PhtB.pro DHYHYIPKN PhtB.pro GHYHYIPKS PhtE.PRO ANPAQPRLSE Majority 250 ANPAQPRLSE PhtD.PRO ANPAQPRLSE PhtB.pro PSVSNPGTTN PhtB.pro BTASD NN PhtE.PRO LVFDPAQITS Majority 100 LIFDPAQITS PhtD.PRO
148 RAD NAVAARAQGRYTTD DGYIPNASDIIEDTGDAYIVPHO 150 RNDGAVALLARSQGRYTTD DGYIPNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTND GYVPNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN PhtB.pro DHYHYIPKN PhtB.pro GHYHYIPKS PhtE.PRO ANPAOPRLSE Majority 250 ANPAOPRLSE PhtD.PRO ANPAOPRLSE PhtB.pro PSVSNPGTTN PhtA.PRO STASD NN PhtE.PRO LVFDPAOITS Majority 100 LIFDPAOITS PhtD.PRO
148 RAD NAVAARA Q GRYTTD D GYIPNASD IIEDTGDAYIVPH 150 RND GAVALARS Q GRYTTD D GYIPNASD IIEDTGDAYIVPH 146 KVN SNVAVARS Q GRYTTND GYVPNPAD IIEDTGNAYIVPH 146 KVN SNVAVARS Q GRYTTND GYPNASD IIEDTGDAYIVPH 146 KVN SNVAVARS Q GRYTTND GYPNASD GYPNASD IIEDTGDAYIVPH 146 KVN SNVAVARS Q GRYTTND GYPNASD IIEDTGNASD GYPNASD IIEDTGNASD GYPNASD IIEDTGNASD GARANT GARAN	DHYHYIPKN Phta.pro DHYHYIPKN Phta.pro GHYHYIPKS Phte.pro ANPAQPRLSE Majority 250 ANPAQPRLSE PhtD.PRO ANPAQPRLSE PhtB.pro PSVSNPGTTN Phta.pro STASD NN Phte.pro LVFDPAQITS PhtD.PRO LIFDPAQITS Phtb.pro LIFDPAQITS Phtb.pro LVFDPAQITS Phtb.pro
148 RADNAVAAARAQGRYTTDDGYIFNASDIIEDTGDAYIVPHO 150 RNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVPHO 146 KVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHO ELSASELAAAEAYLNGK	DHYHYIPKN Phta.pro DHYHYIPKS Phta.pro GHYHYIPKS Phta.pro ANPAQPRLSE Majority 250 ANPAQPRLSE Phtb.pro ANPAQPRLSE Phtb.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTS Majority 300 LIFDPAQITS Phtb.pro LUFDPAQITS Phtb.pro LUFDPAQITS Phtb.pro LVFDPAQITS Phtb.pro LVFDPAQITS Phtb.pro
148 RAD NA VAAARAQ GRYTTDDGY I PNASD I I EDTGDAY I V PHO 150 RNDGAVALARSQ GRYTTDDGY I FNASD I I EDTGDAY I V PHO 146 X V N S N V A V A R S Q GRYTTDDGY V P N P A D I I EDTGNAY I V PHO ELSASELAAAEAY W N GK	DHYHYIPKN PhtB.pro DHYHYIPKS PhtL.PRO GHYHYIPKS PhtL.PRO ANPAOPRLSE Majority 250 ANPAOPRLSE PhtD.PRO ANPAOPRLSE PhtB.pro BYSNPGTTN PhtA.PRO BTASD NN PhtL.PRO LVFDPAOITS Majority 300 LIFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro LVFDPAOITS PhtB.pro
148 RAD NAVAAARA Q G R Y T T D D G Y I P N A S D I I E D T G D A Y I V P H C 150 R N D G A V A L A R S Q G R Y T T D D G Y I F N A S D I I E D T G D A Y I V P H C 146 K V N S N V A V A R S Q G R Y T T N D G Y V P N P A D I I E D T G N A Y I V P H C E L S A S E L A A A E A Y L N G K Q G S R P S S S S S Y N A 199 E L S A S E L A A A E A Y W N G K Q G S R P S S S S S S Y N A 210 220 230 240 198 E L S A S E L A A A E A Y W N G K Q G S R P S S S S S S Y N A 200 E L S A S E L A A A E A Y L S G R G N L S N S R T Y R Q N S D N T S R T N W V I 196 D L S A S E L A A A E A Y L S G R G N L S N S R T Y R Q N S D N T S R T N W V I 250 270 280 290 250 T N L T V T P T Y H Q A N Q G E N I S S L L K E L Y A K P L S E R H V E S D G 251 N H N L T V T P T Y H Q - N Q G E N I S S L L R E L Y A K P L S E R H V E S D G 252 T N H N L T V T P T Y H Q - N Q G E N I S S L L R E L Y A K P L S E R H V E S D G 253 T N T S N N S N T N S Q A S Q S N D I D S L L K Q L Y K L P L S Q R H V E S D G 254 T A R G V A V P H G D H Y H F I P Y S Q M S E L E R I A R I I P L R Y R S N 310 320 330 340	DHYHYIPKN Phta.pro DHYHYIPKS Phta.pro GHYHYIPKS Phta.pro ANPAQPRLSE Majority 250 ANPAQPRLSE Phta.pro ANPAQPRLSE Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTN Phta.pro BYSVSNPGTTS Majority 300 LIFDPAQITS Phta.pro LVPDPAKIIS Phta.pro LVPDPAKIIS Phta.pro LVPDPAKIIS Phta.pro HWVPDSRPEQ Majority
148 RAD NAVAARA Q GRYTTD D GYIPNASD IIEDTG DAYIVPH C 150 RND GAVALARS Q GRYTTD D GYIPNASD IIEDTG D AYIVPH C 146 KVN S NVAVARS Q GRYTTND GYVPNPAD IIEDTG NAYIVPH C ELSASELAAAEAYLNGK	DHYHYIPKN Phts.pro DHYHYIPKS Phts.pro GHYHYIPKS Phts.pro 250 ANPAQPRLSE Phts.pro ANPAQPRLSE Phts.pro ANPAQPRLSE Phts.pro BYSVSNPGTTN Phts.pro BYSVSNPGTTN Phts.pro BYSVSNPGTTN Phts.pro BYSVSNPGTTN Phts.pro BYSVSNPGTTS Majority 300 LIFDPAQITS Phts.pro LVFDPAQITS Phts.pro LVFDPAQITS Phts.pro LVFDPARIIS Phts.pro BVFDPARIIS Phts.pro HWVPDSRPEO Majority 350 HWVPDSRPEO Phts.pro
148 RADNA VAAARA Q GRYTTD D GYIPNASDIIEDTG DAYIVPH G 150 RN D GAVALARS Q GRYTTD D GYIPNASDIIEDTG D AYIVPH G 146 KVN S NVAVARS Q GRYTTND GYVPNPAD IIEDTG NAYIVPH G ELSASELAAAEAYLNGK	DHYHYIPKN Phta.pro DHYHYIPKS Phta.pro GHYHYIPKS Phta.pro ANPAOPRLSE Majority 250 ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro BYSVSNPGTTN Phta.pro BYASD NN Phta.pro BYASD NN Phta.pro LVFDPAOITS Majority LIFDPAOITS Phtb.pro LVFDPAKIIS Phta.pro LVFDPAKIIS Phta.pro LVFDPAKIIS Phta.pro HWVPDSRPEO Majority 350 HWVPDSRPEO Phtb.pro HWVPDSRPEO Phtb.pro
148 RAD NAVAARA Q GRYTTD D GYIPNASD IIEDTG DAYIVPH C 150 RND GAVALARS Q GRYTTD D GYIPNASD IIEDTG D AYIVPH C 146 KVN S NVAVARS Q GRYTTND GYVPNPAD IIEDTG NAYIVPH C ELSASELAAAEAYLNGK	DHYHYIPKN Phta.pro DHYHYIPKS Phta.pro GHYHYIPKS Phta.pro ANPAOPRLSE Majority 250 ANPAOPRLSE Phtb.pro ANPAOPRLSE Phtb.pro BYSVSNPGTTN Phta.pro BYASD NN Phta.pro BYASD NN Phta.pro LVFDPAOITS Majority LIFDPAOITS Phtb.pro LVFDPAKIIS Phta.pro LVFDPAKIIS Phta.pro LVFDPAKIIS Phta.pro HWVPDSRPEO Majority 350 HWVPDSRPEO Phtb.pro HWVPDSRPEO Phtb.pro

6/17

Figure 6(b)

PSPQPTPEPSPSPQPA	PNAPSNP	IDXKLVKEAV	RKVGDGYVFE	ENGV Hajority
360	370	380	390	400
337 PSPQSTPEPSPSPQPA	PNPQPAPSNP	IDEKLVKEAV	RKVGDGYVFE	E-N G V PhED. PRO
	·	IDGKLVKEAV	RKVGDGYVFE	ENG V PheB.pro
JSO PSPOPTPEPSPGPOPA	PNLKIDS	- N S SIL VIS Q LV	RKVGJEJGYVFE	EKGI PhtA.PRO
SRYVPAKDLSAETAAG	LDSKLAKQES	LSHKLGAKKT	DLPSSDREFY	N K A Y Hajority
410	420	430	440	450
3 1 1	IDSKLAKQES		DLPSSDREFY	N K A Y PhtD.PRO
380 S R Y I P A K D L S A E T A A GI 396 S R Y V F A K D L P S E T V K N	IDSKLAKQES Urskuskoesi	LSHKLGITKKT	D L P S S D R E F Y N V A P R D Q E F Y	N K A Y PhtB.pro
J18 G S T V S T N A K P N E V V S S	1 1		SLTTS	DKAY Phta.PRO
			- -	
DLLARIHODLLDNKGR	QVBPEALDNL	LEKLKDVSSD	KVKLVDDILA	F L A P Majority
460	470	480	490	500
437 DLLARIHODLLDNKGR				F L A P PhtD.PRO
	OVDFEALDNL NSDFQALDKL		K V K L VED I L A KER L V D DLLL A	, -
348		KELSS-		Phts. PRO
IRHPERLGKPNAQITY	TDDEIQVAKL		IPDPRDITSD	EGDA Majority
510	520	530	540	550
	* n n w t n W & # f. :	GTVTTTTT	T = D D B D T M C P	
480 IRHPERLGKPNAQITY	TODEIQVARL	AGKYTTEDGY AGKYTAEDGY	IPDPRDITSI	EGDA Phen.pro
1	TODEIQVARL	AGKYT ABDGY	IPDPRDITSE IPDBHDIISE	EGDA PhtB.pro
480 IRHPERLGKPNAQITY	TODEIQVARL	AGKYT ABDGY	IPDPRDITSI	EGDA PhtB.pro
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY	TEDEVRIAGE	AGKYTABDGY ADKYTTSDGY	IPDPRDITSI IPDBHDIISI IPNPKDIVEE	EGDA PhtB.pro EGDA PhtA.PRO ETATA PhtE.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S Q I E Y 353	TEDEVRIAGE	AGKYTABDGY ADKYTTSDGY	IPDPRDITSI IPDBHDIISI IPNPKDIVEE	EGDA PhtB.pro EGDA PhtA.PRO ETATA PhtE.PRO
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY 353	T D D E I O V A K L TED EV R I A Q L L S B A B R A A A Q	AGKYTABDGY ADKYTTSDGY ASDGY AYAKZKGLTP 580	IPDPRDITSI IPDBHDIISI IPNPKDIVEE PSTDHODSG-X	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSOIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS 530 YVTPHHTHSHWIKKDS	T D D E I O V A K L TED EV R I A Q L S E A E R A A A Q S T O L S E A E R A A A Q L S E A E R A A A Q	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP 580 AYAKEKGLTP	IPDPRDITSE IPDBHDIISE IPNPKDIVEE PSTDHODSGN PSTDHODSGN	PEGDA PhtB.pro PEGDA PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS 530 YVTPHHTHSHWIKKDS 546 YVTPHHGHSHWIKKDS	T D D E I O V A K L T E D E V R I A Q L S T O C	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP S80 AYAKEKGLTP AYAKEKGLTP	IPDPRDITSI IPDBHDIISI IPNPKDIVEE PSTDHODSGN PSTDHODSGN PSTDHODSGN	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSOIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS 530 YVTPHHTHSHWIKKDS	T D D E I O V A K L T E D E V R I A Q L S T O C	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP 580 AYAKEKGLTP	IPDPRDITSI IPDBHDIISI IPNPKDIVEE PSTDHODSGN PSTDHODSGN PSTDHODSGN	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS 530 YVTPHHTHSHWIKKDS 546 YVTPHHGHSHWIKKDS	T D D E I O V A K L TED EV R I A Q L 570 L S E A E R A A A Q L S E A E R A A A Q L S E A E R A A A Q L S D K E K V A A Q S N Q I G Q P	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP 580 AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP	IPDPRDITS DIPDED IN SET DE NO DE GENERAL DE SET DE	PEGDA PhtB.pro PEGDA PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro PTGD PhtA.PRO SHEK PhtE.PRO
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSOIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHMTHSHWIKKDS 530 YVTPHMTHSHWIKKDS 546 YVTPHMGHSHWIKKDS 372 YIVRHGDHPHYIPK	T D D E I O V A K L TED EV R I A Q L 570 L S E A E R A A A Q L S E A E R A A A Q L S E A E R A A A Q L S D K E K V A A Q S N Q I G Q P	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP 580 AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP	IPDPRDITS DIPDED IN SET DE NO DE GENERAL DE SET DE	PEGDA PhtB.pro PEGDA PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro PTGD PhtA.PRO SHEK PhtE.PRO
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY 353 YVTPHHTHSHWIKKDS 560 537 YVTPHHTHSHWIKKDS 530 YVTPHHTHSHWIKKDS 546 YVTPHHGHSHWIKKDS 372 YIVRHGDHPHYIPK	T D D E I Q V A K L TED EV R I A Q L S70 L S E A E R A A A Q L S D K E K V A A Q L S D K E K V A A Q C S N Q I G Q P L D R M P Y N L Q Y	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP S80 AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP AYAKEKGLTP	IPDPRDITSI IPDBHDIISI IPNPKDIVEE PSTDHODSGN PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT	EGDA PhtB.pro EGDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro PTGD PhtA.PRO SHEK PhtE.PRO KFEW Majority
480 IRHPERLOKPNAQITY 496 ITHPERLOKPNSQIEY 353	T D D E I Q V A K L T E D E V R I A Q L S E A E R A A A Q S E A E R A A A Q L S E A E R A A A Q L S D K E K V A A Q L S D K E K V A A Q C D R M P Y N L Q Y L D R M P Y N L Q Y	AGKYTABDGY ADKYTTSDGY AVAKEKGLTP S80 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVEE PSTDHODSGN PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI 1PHYDHYHNI IPHYDHYHNI	EGDA PhtB.pro PEGDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtD.PRO TEAK PhtB.pro PTGD PhtA.PRO SHEK PhtE.PRO KFEW Majority 650 KFEW PhtD.PRO KFEW PhtD.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S Q I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 372 Y I V R H G D H P H Y I P K G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K G E K R I P	T D D E I Q V A K L T E D E V R I A Q L S E A E R A A A Q S T O L S E A E R A A A Q L S D K E K V A A Q L S D	AGKYTABDGY ADKYTTSDGY AVAKEKGLTP S80 AYAKEKGLTP	IPDPRDITSI IPDBHDIISI IPNPRDIVEE PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI	EGDA PhtB.pro PtGDA PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtB.pro TEAK PhtB.pro PTGD PhtA.PRO SHEK PhtE.PRO KFEW Majority 650 KFEW PhtB.pro KFEW PhtB.pro
480 IRHPERLGKPNAQITY 496 ITHPERLGKPNSQIEY 353	T D D E I O V A K L T E D E V R I A Q L S E A E R A A A Q S E A E R A A A Q L S E A E R A A A Q L S D K E K V	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP S80 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVKE PSTDHODSGN PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI	EGDA PhtB.pro PtA.PRO PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtB.pro Pt GD PhtA.PRO SHEK PhtE.PRO KFEW PhtB.pro KFEW PhtB.pro KFEW PhtB.pro KFAW PhtA.PRO RFAW PhtA.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S Q I E Y 353 Y V T P H H T H S H W I K K D S 560 537 Y V T P H H T H S H W I K K D S 530 Y V T P H H T H S H W I K K D S 546 Y V T P H H G H S H W I G K D S 372 Y I V R H G D H P H Y I P K G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K G E K R I P	T D D E I O V A K L T E D E V R I A Q L S E A E R A A A Q S E A E R A A A Q L S E A E R A A A Q L S D K E K V	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP S80 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVKE PSTDHODSGN PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI	EGDA PhtB.pro PtA.PRO PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtB.pro Pt GD PhtA.PRO SHEK PhtE.PRO KFEW PhtB.pro KFEW PhtB.pro KFEW PhtB.pro KFAW PhtA.PRO RFAW PhtA.PRO
480 IRHPERLGKPNAQITY 496 ITHPERLGKPNSQIEY 353	T D D E I O V A K L T E D E V R I A Q L S E A E R A A A Q S E A E R A A A Q L S E A E R A A A Q L S D K E K V	AGKYTABDGY ADKYTTSDGY AYAKEKGLTP S80 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVKE PSTDHODSGN PSTDHODSGN PSTDHODSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI	EGDA PhtB.pro PtA.PRO PhtA.PRO TATA PhtZ.PRO TEAK Majority 600 TEAK PhtB.pro Pt GD PhtA.PRO SHEK PhtE.PRO KFEW PhtB.pro KFEW PhtB.pro KFEW PhtB.pro KFAW PhtA.PRO RFAW PhtA.PRO
480 I R H P E R L G K P N A Q I T Y 496 ITH P E R L G K P N S Q I E Y 353 Y V T P H M T H S H W I K K D S 560 537 Y V T P H M T H S H W I K K D S 546 Y V T P H M G H S H W I G K D S 372 Y I V R H G D H P H Y I P K G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K A A K K V P 596 S A A A I Y N R V K G E K R I P 417 H E E D G Y G P D E G L Y E A P K G Y T L E D 660 637 P D E G L Y E A P K G Y T L E D	T D D E I O V A K L T E D E V R I A Q L 570 L S E A E R A A A Q L S E A E R A A A Q L S E A E R A A A Q L S D K E K V A A Q L S D	GKYTABDGY ADKYTTSDGY AVAKEKGLTP 580 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVEE PSTDHQDSGN PSTDHQDSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI GFGNASDHV 690	EGDA PhtB.pro Pt GDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtB.pro PT GD PhtA.PRO SHEK PhtE.PRO KFEW Majority 650 KFEW PhtB.pro KFAW PhtB.pro KFAW PhtB.pro KFAW PhtA.PR TO PhtA.PRO KFAW PhtB.pro KFAW PhtA.PR TO PhtB.PRO KKAW PhtA.PR TO PhtB.PRO KKAW PhtA.PR TO PhtB.PRO
480 I R H P E R L G K P N A Q I T Y 496 I T H P E R L G K P N S O I E Y 353 Y V T P H M T H S H W I K K D S 560 537 Y V T P H M T H S H W I K K D S 546 Y V T P H M G H S H W I G K D S 372 Y I V R H G D H P H Y I P K G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 580 G A E A I Y N R V K A A K K V P 596 S A A A I Y N R V K G E K R I P 417 H E E D G Y G P D E G L Y E A P K G Y T L E D 660	T D D E I O V A K L T E D E V R I A Q L 570 L S E A E R A A A Q L S E A E R A A A Q L S E A E R A A A Q L S D K E K V A A Q L S D	GKYTABDGY ADKYTTSDGY AVAKEKGLTP 580 AYAKEKGLTP	IPDPRDITS IPDBHDIIS IPNPKDIVEE PSTDHQDSGN PSTDHQDSGN PSPDADVKAN SPSLPINPGT IPHYDHYHNI IPHYDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI IPHKDHYHNI GFGNASDHV 690	EGDA PhtB.pro Pt GDA PhtA.PRO TATA PhtE.PRO TEAK Majority 600 TEAK PhtB.pro PT GD PhtA.PRO SHEK PhtE.PRO KFEW Majority 650 KFEW PhtB.pro KFAW PhtB.pro KFAW PhtB.pro KFAW PhtA.PR TO PhtA.PRO KFAW PhtB.pro KFAW PhtA.PR TO PhtB.PRO KKAW PhtA.PR TO PhtB.PRO KKAW PhtA.PR TO PhtB.PRO

7/17

Figure 6(c)

XOXOXNXXXX	XXXEEX -	PE	<u></u>		Hajor
71	10	720 .	730	740	750
				ADNLYKPSTD	•
	KPSEEKPQT	EKPEET	PR1	EEKPQSEKPE	
S H S E D P N K N P L					PhtA.
-					- 22-20
EEXEETPXE	XEXPOVETE	KVEAKLXEA	EXLLXKVT	DPSIKXNAXE	T-L T G L Hajor
70	60	770	780	790	ВОО
				PSIRONAME	TLTGL PhtD.
PEPPESPEP	L		1 1 1 7		
	PEVPOVETE	KVENOLKEA		DSSLKANATE	
)		- PKKDLITEE	Q I		PhtE.
KNNLLLGTXI	ONNTINATA	EKLLALLKE	SXPXXX	- K K	Major
81	10	820	830	840	
KSSLLLGTKI	N N T I S A E V	DSLLALLKE	SQPAPI		PhtD.
KNNLLFGTQI		EKLLALLKE	1 i 		PhtB.
RNNLTLQIMI	DNNISIMARA	EKLLALLKIG	ZIMBIA 2 A 2 ¥ I	1111	· PhtA.
)				-KVRKNI	PhtE.

Decoration 'Decoration #2': Box residues that match the Consensus exactly.

Figure 7(a)

	TCCTATGAGCTTGGA	JTTATCAAGC*	TGGTCAGGT	T	CCTAA Majority
	10	26	30	40	50
61	TCTTACGAGTTGGGAC	TGTATCAAGC			
1	TCCTATGAGCTTGGAG			TAAGAAAGAG1	
1	TCCTATGAACTTGGTC	GTCACCAAGC			
64	GCCTATGCACTAAACC	AGCATC GT	TCG-CAGGA	AAATAAGGACI	ATAA phtE.SEQ
	TCGTGTTTCTTATATA	GATGGTGATE.	AGGCTGGTC	AAAAGGCAGAI	A A C T Majority
	60	70	80	90	100
108	TCGTGTTTCCTATATA	GATGGAAAAC	AAGCGACGC		AATT phtA.SEQ
51	TCGAGTTGCTTATATA	GATGGTGATC		,	
51 111	TCGAGTTTCTTATATA	GATGGTGATC			
111	i cololici ci i xiolo	GATGGCAGCC	A G T C A A G T C J	. U A A A A G T G A A	AACT pate.seq
	TGACACCAGATGAGGT	TAGTAAGAGG	GAGGGGATC	. A C G C T G A G C A	AATT Hajority
	110	120	130	140	150
150					
158 101	T G A C T C C T G A T G A G G T	CAGTAAGAGG			
101	TGACACCAGATGAAGT				· ·
161	TGACACCAGACCAGGT				_
	GTCATCAAGATTACGG	XTCXXGGTTX	TGTGACCTCT	CATEGAGACO	ATTA Majority
	160	170	180	190	200
208	GTCATCAAGATAACAG				
151	GTTATCAAGATTACGG	ATCAAGGTTA:	TGTGACCTCT	CATGGAGACO	ATTA phtB.seq
151 151	G T T A T C A A G A T T A C G G G T C A T C A A G A T T A C G G	A T C A A G G T T A :	T G T G A C C T C 1 T G T G A C C T C 1	C A T G G A G A C C C A T G G A G A C C	ATTA phtB.seq ATTA phtD.SEQ
151	GTTATCAAGATTACGG	A T C A A G G T T A :	T G T G A C C T C 1 T G T G A C C T C 1	C A T G G A G A C C C A T G G A G A C C	ATTA phtB.seq ATTA phtD.SEQ
151 151	G T T A T C A A G A T T A C G G G T C A T C A A G A T T A C G G	ATC	T G T G A C C T C 1 T G T G A C C T C 1 T G T A A C G T C J	C A T G G A G A C C C A T G G A G A C C A C A C G G T G A C C	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ
151 151	G T T A T C A A G A T T A C G G G T C A T C A A G A T T A C G G G T A A T C A A A A T T A C A G	ATC	T G T G A C C T C 1 T G T G A C C T C 1 T G T A A C G T C J	C A T G G A G A C C C A T G G A G A C C A C A C G G T G A C C	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ
151 151	GTTATCAAGATTACGGGTCATCAAGATTACGGGTAATCAAAATTACAGGC	ATCAAGGTTA: ATCAGGGCTA: ATCAGGGCTA: AAGGTTCCTT:	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C 3 A T G A T G C C A 1	CATGGAGACC CATGGAGACC CACGGTGACC	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250
151 151 211	GTTATCAAGATTACGG GTCATCAAGATTACGG GTAATCAAAATTACAG TCATTACTATAATGGC	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: 220 AAGGTTCCTT:	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C 3 A T G A T G C C A 7 230 A T G A C G C T A 7	CATGGAGACC CATGGAGACC CATGAGTGAC 240 CATCAGTGA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ
151 151 211 258 201 201	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C 210 T C A T T A T T A C A A T G G T T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT:	T G T G A C C T C 7 T G T A A C G T C 7 A T G A T G C C A 7 A T G A C G C T A 7 A T G A T G C C A 7 A T G A T G C C A 7	C A T G G A G A C C C A T G G A G A C C C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ
151 151 211 258 201	G T T A T C A A G A T T A C G G G T C A T C A A G A T T A C G G G T A A T C A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A T T A C A A T G G T T C A T T A C T A T A A T G G C	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT:	T G T G A C C T C 7 T G T A A C G T C 7 A T G A T G C C A 7 A T G A C G C T A 7 A T G A T G C C A 7 A T G A T G C C A 7	C A T G G A G A C C C A T G G A G A C C C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A C A T C A G T G A A	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ
151 151 211 258 201 201	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C 210 T C A T T A T T A C A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT:	TGTGACCTCT TGTGACCTCT TGTAACGTCJ ATGATGCCAT ATGACGCTAT ATGATGCCAT ATGATGCCAT	CATGGAGACC CATGGAGACC CATGAGTGAC CATCAGTGAC CATCAGTGAC CATCAGTGAC CATCAGTGAC CATCAGTGAC CATCAGTGAC	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.seq AGAGC phtB.seq AGAAC phtE.SEQ
151 151 211 258 201 201	GTTATCAAGATTACGGGGTCATCAAAAATTACGGGTCATTACTATAATGGCTCATTACTATAATGGCTCATTACTATAATGGCTCATTACTACTATAATGGCTCATTACTATAATGGCTCATTACTATAATGGCTCATTACTATAATGGCTCATTACTATAATGGC	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT:	TGTGACCTCT TGTGACCTCT TGTAACGTCJ ATGATGCCAT ATGATGCCAT ATGATGCCAT ATGATGCCAT ATGATGCCAT	CATGGAGACC CATGGAGACC CATGAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ AGAAC phtE.SEQ
151 151 211 258 201 201 261	GTTATCAAGATTACGGGGTCATCAAAAATTACGGGTCATCAAAAATTACAGGCTCATTACTATAATGGCTTCATTACTATAATGGCTTCATTACTATAATGGCTTCATTACTATAATGGCTTCATTACTATAATGGCTTCATTACTATAATGGCTTCATTACTATAATGGC	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT: AAAGTTCCTT:	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 230 A T G A C G C T A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C 7	CATGGAGACC CATGGAGACC CATGAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG CATCAGTGAG	250 A T T A phtD.SEQ A C T A phtE.SEQ A G A A T phtA.SEQ A G A G C phtB.seq A G A G C phtB.seq A G A A C phtE.SEQ A G A A C phtE.SEQ A C A A T Majority 300
151 151 211 258 201 201 261	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G G T C A T T A C T A T A A T G G G T C C T C A T G A A A G A T C C T A C T C A T G A A A G A T C C	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT:	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 230 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C A 7 A T G A T G C C C C 7 280 C T A A A A G A T C	CATGGAGACC CATGGAGACC CATGAGTGAG CATCAGTGAG CAGATATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtB.seq AGAAC phtE.SEQ AGAAC phtE.SEQ ATTAAT phtA.SEQ
151 151 211 258 201 201 261	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G G T C A T T A C T A T A A T G G G T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C	ATCAAGGTTA: ATCAAGGTTA: ATCAGGGCTA: AAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT:	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C 7 T T G A A G G A T 7 T T G A A G G A T 7	240 CATGGAGACA CATGGAGACA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ AGAAT phtA.SEQ AGAAT phtA.SEQ AGAAT phtB.seq CAAT phtA.SEQ CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G G T C A T T A C T A T A A T G G G T C C T C A T G A A A G A T C C T A C T C A T G A A A G A T C C	ATCAAGGTTA; ATCAAGGTTA; ATCAGGGCTA; 220 AAGGTTCCTT; AAGGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; GAATTATCAG; GAATTATCAG; GAATTATCAG;	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C A 7 A T G A A G G A T 7 T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7	240 CATCAGTGAA 240 CATCAGTGAA CCAGACATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ AGAGC Majority 250 AGAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ AAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C	ATCAAGGTTA: ATCAAGGTTA: ATCAAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAACTATAAGG	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A A G G A T 7 T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G G T C A T T A C T A T A A T G G G T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C T T G A T G A A G G T G G T T C C T C A T G A A G G T G G T T C C T C A T G A A G G T G G T T	ATCAAGGTTA; ATCAAGGTTA; ATCAAGGTTCCTT; AAGGTTCCTT; AAGGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAACTATAAG; GAATTATCAG; GAATTATCAG; AAACTATATCAG; AAACTATATCAG;	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A A G G A T 7 T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C	ATCAAGGTTA: ATCAAGGTTA: ATCAAGGTTCCTT: AAGGTTCCTT: AAGGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAAGTTCCTT: AAACTATAAGG	T G T G A C C T C 7 T G T G A C C T C 7 T G T A A C G T C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A A G G A T 7 T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7 T T G A A G G A T 7	CATGGAGACC CATGGAGACC CATCAGTGAA 240 CATCAGTGAA CATTGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ AGAGC phtB.seq AGAGC phtD.SEQ AGAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251 311	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C 210 T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C T T G A T G A A G G T G G T T 310 G A G G T C A A G G G T G G A T	ATCAAGGTTA; ATCAAGGTTA; ATCAAGGTTCCTT; AAGGTTCCTT; AAGGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAACTATAAG; AAACTATATCAG; AAACTATATCAG; AAACTATATCAA; ATGTTATCAA; 320	T G T G A C C T C 7 T G T A A C G T C 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C 7 A T G A A G G A T 7 T G A A G A T G G 7 T G G T A G A T G G 7	240 CATGGAGACA CATGAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CCATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATCAGTGAA CATATTGT CAGACATTGT CAAAATACTATCGT	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq
151 151 211 258 201 201 261 308 251 251 311	G T T A T C A A G A T T A C G G G T C A T C A A G A T T A C G G G T A A T C A A A A T T A C A G T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G G T C A T T A C T A T A A T G G G T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C T T G A T G A A A G A T C C G A A G T C A A G G G T G G T T 310 G A G G T C A A G G G T G G T T G A A A G G T C A A G G G T G G T T	ATCAAGGTTA; ATCAAGGTTA; ATCAAGGCTA; 220 AAGGTTCCTT; AAAGGTTCCTT; AAAGTTCCTT; AAAGTTATCAG; AAACTATATCAG; ATGTTATCAAG	T G T G A C C T C 7 T G T A A C G T C 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A A G G A T 7 T G A A A G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G G A T 7 T G A A G A T G G 7 G G T A G A T G G 7 G G T A G A T G G 7	240 CATCAGTGAA 240 CATCAGTGAA CATTTAGTGAA AAATACTATGTAA 340 AAATACTATG	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq TTTA Majority 150 TTTA phtA.SEQ TTTA phtA.SEQ
151 151 211 258 201 201 261 308 251 251 311	G T T A T C A A G A T T A C G G G T C A T C A A A A A T T A C A G T C A T T A C T A T A A T G G C 210 T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C A T T A C T A T A A T G G C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C C T C A T G A A A G A T C C T C T T G A T G A A G G T G G T T 310 G A G G T C A A G G G T G G A T	ATCAAGGTTA; ATCAAGGTTA; ATCAAGGCTA; 220 AAGGTTCCTT; AAGGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAAGTTCCTT; AAACTATAAG; AAACTATATCAA; GAATTATCAA; AAACTATATCAA; AAACTATATCAA; ATGTTATCAA;	T G T G A C C T C 7 T G T A A C G T C 7 A T G A T G C C A 7 A T G A T G C C A 7 A T G A T G C C C 7 A T G A A G G A T 7 T G A A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7 T G G T A G A T G G 7	240 CATCAGTGAA 240 CATCAGTGAA CATTTAGTGAA AAATACTATGTAA AAATACTATATGAA AAATACTATATGAA	ATTA phtB.seq ATTA phtD.SEQ ACTA phtE.SEQ GAGC Majority 250 GAAT phtA.SEQ GAGC phtB.seq GAGC phtD.SEQ GAAC phtE.SEQ CAAT Majority 300 TAAT phtA.SEQ CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq CAAT phtB.seq TTTA Majority 150 TTTA phtA.SEQ TTTA phtB.seq TTTA phtB.seq

Figure 7(b)

	CCTTAAGGATGCAGC	CATGCGGATA	ATGTTCGGAC	******	ATTA Majority
	360	370	380	390	400
408	CCTTAAGGATGCTGC	CACGCGGATA	ACGTCCGTAC	AAAGAGGAA	ATCA Dhta SEO
351		CATGCGGATA			
351	CCTTAAGGATGCAGC	CATGCGGATA	ATATTCGGAC	AAAAGAAGA'G	ATTA phtD.SEQ
411	CCTGAAAGATGCAGC	CATGCTGATA	ATGTTCGAAC	TAAAGATGAA	ATCA phte.SEQ
	ATCGTCAGAAGCAGG	ACATAGTCAT	AATCATGAGG	GTGGAXCT	- A Majority
	410	420	430	440	-
					450
458 401	ATCGACAAAAACAAGA				phtB.seq
401	AACGTCAGAAGCAGG				phtD.SEQ
461	ATCGTCAAAAACAAG	ACATOTCAAA	G A T A A T G A G -		- A A G phtE.SEQ
					0001 Wedender
	GATGATXXTGCTGTTC			CGCTATACAA	C G G A MAJORITY
	460	470	480	490	500
508	AACGATGGTGCTGTTC	CCTTGGCACG			
445	GATAAT GCTGTTC	CTGCAGCCAG			
445 499	A A C G A T C A A G C A G T A C				-
437	U ! ! X X C ! C ! X X ! C ! ! .				C w w w hurenbed
	TGATGGTTATATCTTT	AATGCATCTG	ATATCATTOA	GGATACGGGT	GATG Majority
	510	520	รวุ๋ง	540	550
558	TGATGGTTATATCTT	AATGCTTCTG			
			. T . T C . T . G .	GGATACTGGT	GATG phta.SEO
492	TGATGGGTATATCTT	AATGCATCTG.	A T A T C A T T G A	GGACACGGGT	GATG phtB.seo
495	TGATGGTTATATCTT			. G G A C A C G G G T	GATG phtB.seq GATG phtD.SEQ
				. G G A C A C G G G T	GATG phtB.seq GATG phtD.SEQ
495	TGATGGTTATATCTT		A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A	. G G A C A C G G G T . G G A C A C G G G T . A G A T A C G G G T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ
495	TGATGGTTATATCTTC		A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A	. G G A C A C G G G T . G G A C A C G G G T . A G A T A C G G G T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ
495 549	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560	AATGCATCTG AATGCATCTG AATCCAGCTG TGGCGATCAT	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A 590	GATG phtB.seo GATG phtD.SEQ AATG phtE.SEQ TGAG Majority
495	TGATGGTTATATCTTC TGATGGTTATGTCTTT	AATGCATCTG AATGCATCTG AATCCAGCTG TGGCGATCAT	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A 580 F A C C A T T A C A	GGACACGGGT GGACACGGGT AGATACGGGT TTCCTAAGAA 590	GATG phtB.seo GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEO
495 549 608 542 545	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	**************************************	X T X T C X T T G X X T X T C X T T G X X T X T T X T C G X	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtD.SEQ
495 549 608 542	TGATGGTTATATCTTC TGATGGTTATATCTTTC CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC	**************************************	X T X T C X T T G X X T X T C X T T G X X T X T T X T C G X	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtD.SEQ
495 549 608 542 545	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	AATGCATCTG AATGCATCTG AATCCAGCTG TGGCGATCAT 570 TGGAGATCAT CGGCGACCAT CGGCGACCAT	A T A T C A T T G A A T A T C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T 590 T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A	GATG phtB.seq AATG phtB.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seq TGAG phtD.SEQ CGAT phtE.SEQ
495 549 608 542 545	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	AATGCATCTG AATGCATCTG AATCCAGCTG TGGCGATCAT 570 TGGAGATCAT CGGCGACCAT CGGCGACCAT	T	GGACACGGGT GGACACGGGT AGAACACGGGT TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA TTCCTAAGAA	GATG phtB.seq AATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seq TGAG phtD.SEQ CGAT phtE.SEQ
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG 570 TGGAGATCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A T C A C T A C A	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A	GATG phtB.seq AATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq CGAT phtE.SEQ GGGA Majority
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG TGGAGATCAT CGGCGACCAT CGCCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGCCGACCAT CGCCGACCAT CCCCACCAT CCCCCACCAT CCCCCCCCCC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A T C A C T A C A A G A A G C C 630 A G A A G C C T T C	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A T T G G A T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq GGAT phtE.SEQ GGGA Majority 650 GAG phtA.SEQ
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG	AATGCATCTG AATGCATCTG AATGCAGCTG TGGAGATCAT S70 TGGAGATCAT CGGCGACCAT CGCCGACCAT CGGCGACCAT CGCGACCAT CGCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CGCCACCAT CCCCACCAT CCCCCACCAT CCCCCACCAT CCCCCACCAT CCCCCACCAT CCCCCACCAT CCCCCCCACCAT CCCCCCCC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A T C A C T A C A A G A A G C C T T C A G A A G C C T T C A G A A G C C T T C A G A A G C C T T C A G A A G C C T T C A G A A G C C T T C	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T 590 T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEO TGAG phtB.seq TGAG phtB.seq CGAT phtE.SEQ GGGA Majority 650 GAGG phtA.SEO GGGA phtB.seq
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAGT TTATCAGCTAGCGAGT TTATCAGCTAGCGAGT	AATGCATCTG AATGCATCTG AATGCATCAT S70 TGGAGATCAT CGGCGACCAT CGGCGACCAT TGGAGGTCAC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C C A A A A G - T A T T T G G A T - T A T T G G A A T	GATG phtB.seq GATG phtD.SEQ AATG phtD.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq GGGA Majority 650 GGGA phtB.seq GGGA phtB.seq GGGA phtB.seq
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATGTCTTC 560 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATATCGTTCCTC TTATATCGTTCCTC TTATCAGCTAGCGAG	AATGCATCTG AATGCATCTG AATGCATCAT 570 TGGAGATCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT CGGCGACCAT TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC TAGCTGCTGC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A T C A C T A C A A G A A G C C T T C A G A A G C C T T C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C A G A A G C C	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A G A A T T C C T A A G A A T T C C T A G A A T T C C T A G A A T T C C C A A A A G - T A T T G G A A T - T A T T G G A A T - T A T T G G C T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG Majority 600 TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq GGAT phtE.SEQ GGGA Majority 650 GAGG phtA.SEQ GGGA phtB.seq GGGA phtB.seq GGGA phtB.seq
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AGCAAAT	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG S70 TGGAGATCAT CGGCGACCAT CGGCACCAT CGGCGACCAT CGGCACCAT CGCCACCAT CCCACCAT CGCCACCAT CCCACCAT CCCACCAT CCCACCAT CCCCACCAT CCCACC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A A G C C A G A A G C	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A G A T T C T A T T G G A T T C T A T T G G A A T T C T A G T T A T	GATG phtB.seq GATG phtD.SEQ AATG phtD.SEQ TGAG PhtA.SEO TGAG phtB.seq TGAG phtB.seq TGAG phtB.seq GGAT phtE.SEQ GGA PhtB.seq GGGA phtB.seq
495 549 608 542 545 599	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AGCAAAT 660	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG S70 TGGAGATCAT CGGCGACCAT CGGCACCAT CGCACCAT CGCCACCAT CGCCACCAT	ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA 580 ACCATTACA ACCATTACA ACACCATTACA ACACCATTACA ACAACCC 630 ACAACCC	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A G A T T C T A T T G G A A T T C T A T T G G A A T T C T A G T T A T 690	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG PhtA.SEQ TGAG phtB.seq TGAG phtB.seq TGAG phtB.seq GGGA Majority 650 GGGA phtB.seq
495 549 608 542 545 599 658 592 595 649	TGATGGTTATATCTTC TGATGGTTATATCTTT CTTATATCGTTCCTC S60 CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC CTTATATCGTTCCTC TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG TTATCAGCTAGCGAG AGCAAAT	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG TGGAGATCAT CGGCGACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCCACCAT CGGCACCAT CGCACCAT CGCACCAT CGGCACCAT CGCACCAT CCCACCAT CCCACCAT CCCACC	A T A T C A T T G A A T A T C A T T G A A T A T T A T C G A F A C C A T T A C A F A C C A T T A C A F A C C A T T A C A F A T C A C T A C A A G A A G C C A G A A G C C F A A A G C C F A A A G C C F A A A G C C F A A A G C A C F A A A G C A C A A A A 680 620 630	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T 590 T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A G A T T A T T G G A T T A T T G G A T T T C T A G T T A T 690 T A G C G A T A A C	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG phtA.SEQ TGAG phtB.seq TGAG phtB.seq TGAG phtB.seq GGAT phtE.SEQ GGAT phtE.SEQ GGAAT phtE.SEQ GGAA phtB.seq GGGAA phtB.seq GGGAA phtB.seq GGGAAA phtB.seq GGGAAA phtB.seq GGGAAAA phtB.seq GGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
495 549 608 542 545 599 658 592 595 649	T G A T G G T T A T A T C T T C T G A T G G T T A T G T C C T C T 560 C T T A T A T C G T T C C T C T C T T A T A T C G T T C C T C T C T T A T A T C G T T C C T C T C T T A T A T C G T T C C T C T C T T A T A T C G T T C C T C T C T T A T A T C G T T C C T C T T T A T C A G C T A G C G A G T T A T C	AATGCATCTG AATGCATCTG AATGCATCTG AATCCAGCTG TGGAGATCAT CGGCGACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCACCAT CGGCCACCAT CGGCACCAT CGCACCAT CGCACCAT CGGCACCAT CGCACCAT CCCACCAT CCCACCAT CCCACC	ATATCATTGA ATATCATTGA ATATCATTGA ATATCATTACA 580 FACCATTACA 630 AGAAGCC CTTCTTCAAG	G G A C A C G G G T G G A C A C G G G T A G A T A C G G G T 590 T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A A G A A T T C C T A G A A T T C C T A G G A T - T A T T G G A A T - A T C T G G C T 690 T A G C G A T A A C T T C T A G T T A T	GATG phtB.seq GATG phtD.SEQ AATG phtE.SEQ TGAG PhtA.SEQ TGAG phtB.seq TGAG phtB.seq TGAG phtB.SEQ GGAT phtE.SEQ GGGA phtB.seq GGGA phtB.seq GGGA phtB.seq GGGA phtB.seq ACTT Majority 700 ACTT phtA.SEQ AATG phtB.seq

10/17

Figure 7(c)

	CAA-ATCCAGCTCAGT	λ C C λ λ	GATTGTCAGAG	AACCACAAT	C T Majority
	710	720	730	740	750
	•	1			
758	CAAGAACAAACTGGGT			A A C T A C A A A T A A A C C A C A A T	
677	CAA-ATCCAGCTCA	YCC Y Y		A A C C A C A A T	-
680	CAA-ATCCAGCTCA		GACAAT AAC		-
728	CAXCAGCI-AGI				•
	GACA-AAGCTGTCACT	CCAACATTAT	CA-TCAAGCAA	ATCAAGGTGAA	A A Majority
	760	770	780	790	800
	1		•		G A phra SEO
808	AACACAAGCAACAACA		C		
717 720				ATCAAGGGGAA	•
759	AGCAAAAG-GATCA	• • • • • •	CAAGCCAGCAA	A T A A A T C T G A A	A A phtE.SEQ
			•		
	CATTTCAAGTCTTTTG	CGTGAATTGT	ATGCTAAACCT	TTATCAGAACG	CC Majority
	810	820	830	840	850
	CATTGATAGTCTCTTG	AAACAGCTCT	ACAAACTGCCT	TTGAGTCAACG	A C phtA.SEQ
858 756	CATTTCAAGCCTTTTA		ATGCTAAACCC		C C phtB.seq
759	CATTTCAAGCCTTTTA	CGTGAATTGT	SOSKARTOREK	TTATCAGAACG	C C phtD.SEQ
801	TCTCCAGAGTCTTTTG	AAGGAACTCT	ATGATTCACCT	Y @ C @ C C C Y Y C @	T T phtE.SEQ
					C) Voiceite
	ATGTGGAATCTGATGG	CCTTGTTTTT	GACCCAGCGCA	AATCACAAGTC	T MAJORICY
	860	870	880	890	900
908	ATGTAGAATCTGATGG	CCTTGTCTTT	GATCCAGCACA	AATCACAAGTC	G A phtA.SEQ
806	ATGTGGAATCTGATGG	CCTTATTTTC		AATCACAAGTC	G A phtB.seq
809	ATGTGGAATCTGATGG		GACCCAGCGCA		
851	ACAGTGAATCAGATGG	CCTGGTCTTT	CACCCTGCTAA	GATTATCAGTC	G T PREE.SEQ
	ACCGCCAGAGGTGTTG		. TGGTGACCATT	ACCACTTTATC	C C Majority
	XEEGECKGXGC1G11G				•
	910	920	930	940	950
958	ACAGCTAGAGGTGTTG	CAGTGCCACA	CGGAGATCATT	ACCACTTCATC	
856	ACCGCCAGAGGTGTAG	CTGTCCCTCA		ACCACTTTATC	C C phtB.seq
859	ACCGCCAGAGGTGTAG	• •		ACCACTTTATC	CC phtD.SEQ
901	ACACCAAATGGAGTTG	CGATTCCGCA	LTGGCGACCATT	XCCXC111X1.	C C p
	TTATGAACAAATGTCT	GAATTGGAAG		CGTATTATTCC	CC Majority
		970	980	990	1000
	960				1_
	TTACTCTCAAATGTCT				CC phtB.seq
906	TTATGAACAAATGTCT		c	• • • • • • •	To the second se
909 951	. T T A T G A A C A A A T G T C T				phtE.SEO
321					
	TTCGTTATCGTTCAAA	CCATTGGGTA	ACCAGATTCAAG	A C C A G A A G A A G	C A Majority
	1010	1020	1030	1040	1050
		<u> </u>		GCCAGAACAAG	C A phta. SEO
956			• • • • • • • • • • • • •	A C C A G A A G A A G	
	TTCGTTATCGTTCAAA				
993		GGTC	G C C T	ATCAGTGGAAG	T G phtE.SEQ

11/17

Figure 7(d)

AGTCCACAATCGACTCCGGAACCTAGTCCA	AGTCCGCAACCTGCACCAAA Majority	Y
1060 1070 108	0 1090 1100	
1108 AGTCCACAACCGACTCCGGAACCTAGTCCA	GGCCCGCAACCTGCACCAAA phtA.SEG	Q
1006 AGTCCACAACCGACTCCAGAACCTAGTCCA	AGTCCGCAACC phtB.sec	Q
1009 AGTCCACAATCGACTCCGGAACCTAGTCCA		_
1013 GTTCTACAGTTTCTACA	AA	đ
TC-T-AAAGCTCCAAGCAATCCAATTGA	TG - GAAATTGGTCAAAGAAG Majority	v
		•
	TTGGTTAGTCAGC phtA.SE(_
1059 TCCTCAACCAGCTCCAAGCAATCCAATTGA	TGAGAAATTGGTCAAAGAAG phtD.SEG	•
1039 C C T A A	TG phtE.SE	_
_	•	
CTGTTCGAAAAGTAGGCGATGGTTATGTCT	TTGAGGAGAATGGAGTTTCT Majority	Y
1150 1170 118	1190 1200	
1196 TGGTACGAAAGTTGGGGAAGGATATGTAT	TCGAAGAAAGGGCATCTCT phta.SE	D.
1088 CTGTTCGAAAAGTAGGCGATGGTTATGTCT	TTGAGGAGAATGGAGTTTCT phtB.sec	•
1109 CTGTTCGAAAAGTAGGCGATGGTTATGTCT	TTGAGGAGAATGGAGTTTCT phtD.SE	_
1046	phts.se	Q
CCTTATATCCCAGCCAAGGATCTTTCAGCA	GAAACAGCAGCAGCATTGA Majorit	Y
1210 1220 123	0 1240 1250	
1246 CGTTATGTCTTTGCGAAAGATTTACCATCT		
1246 CGTTATGTCTTTGCGAAAGATTAACGATCT	GAAACTGTTAAAAATCTTGA phtA.SE	:Q
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA		
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA	GAAACAGCAGCAGGCATTGA phtB.se	Q: Q:
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA	GAAACAGCAGCAGGCATTGA phtB.se	pi Q:
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE 	.Q .Q
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	.Q .Q
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	. A . C . C . C
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	50 50 50 50
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	55 57 58 50 50 50
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	50 50 50 50 50 60
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	50 50 50 50 50 60
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtE.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAAA phtA.SE TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se	20 20 20 20 21 20 20
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtB.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAAA phtA.SE TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAGCTAAGA phtB.se	20 20 20 20 21 20 20
1138 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1159 CGTTATATCCCAGCCAAGGATCTTTCAGCA 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtB.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAAA phtA.SE TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAGCTAAGA phtB.se TCATAAGCTAGGAGCTAAGA phtC.SE	2.0 2.0 2.0 2.0 3.0 3.0 3.0 4.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtB.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAAA phtA.SE TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAGCTAAGA phtB.se TCATAAGCTAGGAGCTAAGA phtC.SE	25 25 25 25 25 25 25 25 25 25 25 25 25 2
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	20 cm
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	20 20 20 20 20 20 20 20 20 20 20 20 20 2
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtB.SE TCATAAGCTAGGAGCTAAGA Majorit 0	25 25 25 25 25 25 25 25 25 25 25 25 25 2
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtB.seAGGC phts.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAAA phtA.SE TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TTATACGATAAGGCTTATGAC Majorit 0 1340 1350 TTTATGATAAAGGCTTATGAC PhtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se	25 25 25 25 25 25 25 25 25 25 25 25 25 2
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtB.seAGGC phts.Se TCATAAGCTAGGAGCTAAGA Hajorit 0	A SOLUTION AND THE SOLU
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	SA S
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SEAGGC phtE.SE TCATAAGCTAGGAGCTAAGA Majorit 0 1290 1300 ACACACTTTAACTGCTAAGA PhtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TCATAAGCTAGGAACTAAGA phtB.se TTTACGATAAGGCTTATGAC Majorit 0 1340 1350 TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGCTTATGAC phtB.se TTTACAATAAGGGTCGACAAGT Majorit 0 1390 1400 GNAAATAAGGGTCGACAAGT phtB.se GATAATAAGGGTCGACAAGT phtB.se	Section Associated Aso
1138 C G T T A T A T C C C A G C C A A G G A T C T T T C A G C A 1159 C G T T A T A T C C C C A G C C A A G G A T C T T T C A G C A 1062	GAAACAGCAGCAGGCATTGA phtB.se GAAACAGCAGCAGGCATTGA phtD.SE	STATE OF THE STATE

12/17

Figure 7(e)

TGATTTTGAGGCTT%	GATAACCTGTTGGAACGAC	TCAAGGATGTCTCAA Hajority
1410	1420 1430	1440 1450
1446 TGATTTCCAAGCCTTA	GACAAATTATTAGAACGCT	TGAATGATGAATCGA phta.SEO
1338 TGATTTTGAGGCTTTG	GATAACCTGTTGGAACGAC	TCAAGGATGTCTCAA phtB.seq
1359 TGATTTGAGGCTTTG	G A T A A C C T G T T G G A A C G A C	TCAAGGATGTCCCAA phtD.SEQ
1107		phtE.SEQ
GTGXTXXXGTCXXGTT		TTCTTAGCTCCGATT Majority
1460	1470 1480	1490 1500
1496 CTAATAAAGAAAATT	GGTAGATGATTATTGGCA	TTCCTAGCACCAATT phtA.SEQ
1388 G T G A T A A A G T C A A G T T	AGTGGAAGATATTCTTGCC	TTCTTAGCTCCGATT phtB.seq
1409 G T G A T A A A G T C A A G T T	AGTGGATGATATTCTTGCC	TTCTTAGCTCCGATT phtD.SEQ
1107		CTCTT phte.seQ
CGTCATCCAGAACGTT	TAGGAAAACCAAATGCGCA	AATTACCTACACTGA Hajority
1510	1520 1530	1540 1550
	TTGGCAAACCAAATTCTCA	
1546 A C C C A T C C A G A G C G A C	T A G G A A A A C C A A A T G C G C A	AATTACCTACACTGA phtB.seq
1459 CGTCATCCAGAACGTT	T A G G A A A A C C A A A T G C G C A	AATTACCTACACTGA phtD.SEQ
1115		phtE.SEC
TGATGAGATTCAAGTA	GCCXXGTTGGCXGGCXXGT	ACACAGCATCAGATG Majority
1560	1570 1580	1590 1600
		ATACAACGTCAGATG phta.SEO
		ACACAGCAGAAGACG phtB.seo
	CCCXXGTTGGCXGGCXXGT	ACACAACAGAAGACG phtD.SEQ
1115		CAGCATCTGATG phtE.SEO
GTTATATTTTGATCC	T C G T G A T A T A A C C A G T G A T	GAGGGGATGCCTAT Majority
1610	1620 1630	1640 1650
1646 GTTACATTTTTGATGA		GAAGGAGATGCATAT phtA.SEQ GAGGGGATGCCTAT phtB.seq
1559 6 T T A T A T C T T T G A T C C		GAGGGGATGCCTAT phtD.SEQ
1127 G T T A T A T T T T A A T C C		phtE.SEQ
GTAACTCCACATATGA	CCCATAGCCACTGGATTAA	AAAAGATAGTTTGTC Majority
1660	1670 1680	1690 1700
1696 GTAACGCCTCATATGG	GCCATAGTCACTGGATTGG	AAAAGATAGCCTTTC phta.SEQ
1588 GTAACTCCACATATGA	. C C C A T A G C C A C T G G A T T A A	AAAAGATAGTTTGTC phtB.seq
1609 GTAACTCCACATATGA	CCCATAGCCACTGGATTAA	AAAAGATAGTTTGTC phtD.SEQ
1143		AAAAGATATC phte.SEQ
TC.).GCTG.).G.).GCG		AAGAGAAAGGTTTGA Hajorit
1710	1720 1730	1740 1750
		AAGAAAAGGTATCC phta.seo
1638 TGAAGCTGAGAGAGCG	G	AAGAGAAAGGTTTGA phtB.seq
1550		1 1 0 1 0 1 1 1 0 0 0 0 0 0 1 -L-0 CTO
		AAGAGAAAGGTTTGA phtD.SEQ

13/17

Figure 7(f)

CCCCTCCTTCGACAG	CATCAGGA	TTCAGGAAAT	ACTGAGGCAA	AAGGA Majority
1760	1770	1780	1790	1800
1796 TACCTCCATCTCCAG	CGCAGATGT	TAAAGCAAAm	•	
1688 C C C C T C C T T C G A C A G	CCATCAGGA		ACTGAGGGAG	ATAGT phtA.SEQ AAGGA phtB.seq
1709 CCCCTCCTTCGACAG	CCATCAGGA	TTCAGGAAAT	ACTGAGGCAA	AAGGA phtD.SEQ
1167				phtE.SEQ
GCAGAAGCTATCTAC	ACCGXGTGA	LAGCAGCTAA	GAAGGTGCCA	C'T T G A Majority
1810	1820	1830	1840	1850
1846 GCAGCAGCTATTAC	ATCGTGTGA	AGGGGAAAA	ACGAATTCCA	CTCGT phtA.SEQ
1738 GCAGAAGCTATCTAC	ACCGHGTGA	LAGCAGCTAA	GAAGGTGCCA	CTTGA phtB.seg
1759 GCAGAAGCTATCTAC	ACCGCGTGAI	LAGCAGCTAA	GAAGGTGCCA	CTTGA pheD.SEQ
1167 TAC	GCT			phtE.SEQ
TCGTATGCCTTACAA1	CTTCAATAT	CTGTAGAAG	TC11111CGG	TAGTT Majority
1860	•		***	
	1870	1880	1890	1900
1896 T C G A C T T C C A T A T A T C				TAATT phta.SEQ
1809 T C G T A T G C C T T A C A A 1	'	CTGTAGAAG		TAGTT phtB.seq
1174	'	CTGTAGAAG TTGTAAGA-	T C A A A A A C G G	TAGTT phtD.SEQ
				phtE.SEQ
TAATCATACCTCATT	TGATCATTAC	CATAACATT	AAATTTGAGT	GGTTT Majority
1910	1920	1930	1940	1950
1946 9 6 3 9 9 3 9 9 6 6 9 6 3 9 3 3				
1946 I GATTATTCCTCATAA		CATAATATT	AAATTTGCTT	GGTTT phta SEC
1946 TGATTATTCCTCATAA 1838 TAATCATACCTCATTA	TGACCATTAC	CATAACATC	AAATTTGAGT	GGTTT phtB.sec
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA	. T	CATAACATC	AAATTTGAGT	GGTTT phtB.sec
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA	TGACCATTAC	CATAACATC	A	GGTTT phtB.sec
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATG	T G A C C A T T A C T G A C C A T T A C T G A T C A T T T C	CATAACATC CATAACATC CATTACATT	A A A T T T G A G T A A A T T T G A G T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG	TGACCATTAC TGACCATTAC TGATCATTTC	CATAACATC CATAACATC CATTACATT	A A A T T T G A G T A A A T T T G A G T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA	CATAACATC CATAACATC CATAACATT GGGGTATAC	A A A T T T G A G T A A A T T T G A G T 	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980	AAATTTGAGT AAATTTGAGT TCTTGAGGAT 1990	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970 AAGCTCCAAA	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	AAATTTGAGT AAATTTGAGT TCTTGAGGAT 1990 CTTGGAAGAT	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG	T G A C C A T T A C T G A C C A T T T C A G G C A C C T A A 1970 A A G C T C C A A A A G G C A C C T A A A G G C A C C T A A	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG	T G A C C A T T A C T G A C C A T T T C A G G C A C C T A A 1970 A A G C T C C A A A A G G C A C C T A A A G G C A C C T A A	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	1990 C T T G A G G A T T C T T G A G G A T C T T G G A A G A T T C T T G A G G A T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC	1990 C T T G A G G A T T C T T G A G G A T C T T G G A A G A T T C T T G A G G A T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A 2040	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186 CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C A C A	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C T A phtA.SEO T T C A G PhtB.seo
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA TATGTCGAAC TATGTCGAAC TATGTCGAAC	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG ATCCAAACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C A G phtB.seq T T C A G phtB.seq T T C A G phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA TATGTCGAAC TATGTCGAAC TATGTCGAAC	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG ATCCAAACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C A G phtB.seq T T C A G phtB.seq T T C A G phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	T G A C C A T T A C T G A C C A T T A C A G G C A C C T A A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G G C A C C T A A G	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG ATCCAAACG ATCCAAACG ATCCAAACG ATCCAAACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A C G T C C G C A	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C A G phtB.seq T T C A A phtE.SEO
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA CAAA CGCTAGGAC CGCTAGCGAC CGCTAGCGAC	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG ATCCAAACG ATCCAAACG ATCCAAACG ATCCAAACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A C G T C C G C A	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtA.SEO C T T T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C A G phtB.seq T T C A A phtE.SEO
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC TGATCATTTC AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA CO200 TACGTAGAAC TATGTCGAAC TATGTCGAAC CAAATTGGGC CGCTAGCGAC	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG 2030 ACCCTGACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A C G T C C G C A A A C G T C C G C A A C G T	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtB.seq C T T T T phtB.seq C T T T T phtB.seq T T C A G Majority 2050 T T C A G phtB.seq T T C A A phtE.SEO A A G A T Majority 2100
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210 GGCGACTGTCAAGTAC 2010 2046 TGCGACGATTAAGTAC 1938 GGCGACTGTCAAGTAC 1959 GGCGACTGTCAAGTAC 1959 GGCGACTGTCAAGTAC 2050 ATAATGGTTTTGGTAA 2060 2096 ATGATGGATGGGGGCAA	TGACCATTAC TGACCATTAC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA COLAAA TATGTCGAAC TATGTCGAAC TATGTCGAAC TATGTCGAAC CAAATTGGGC CGCTAGCGAC 2070 TGCCAGTGAG	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG ATCCAAACG	A A A T T T G A G T A A A T T T G A G T T C T T G A G G A T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A C C G T C C G C A A C C G T C C G C C C C C C C C C C C C C	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtB.seq C T T T T phtB.seq C T T T T phtD.SEO phtE.SEO T T C A G Majority 2050 T T C A G phtB.seq T T C A G phtB.seq T T C A G phtB.seq T T C A G phtD.SEO T C C A A phtE.SEO A A G A T Majority 2100 A A G A C phtA.SEO
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATGG GACGAAGGCCTTTATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210	TGACCATTAC TGACCATTAC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA TATGTCGAAC TATGTCGAAC TATGTCGAAC TATGTCGAAC TATGTCGAAC CAAATTGGGC CGCTAGCGAC CGCTAGCGAC CGCTAGCGAC CGCTAGCGAC	CATAACATC CATAACATC CATAACATT GGGGTATAC GGGGTATAC GGGGTATAC GGGGTATAC ATCCAGACG ATCCAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG ATCCAAAACG	A A A T T T G A G T A A A T T T G A G T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C A A G A A G A A A C A A G A	GGTTT phtB.seq GGTTT phtD.SEC phtE.SEC CTTTT Majority 2000 TTGTT phtB.seq CTTTT phtB.seq CTTTT phtB.seq CTTTT phtB.seq TTCAG Majority 2050 TTCAG phtB.seq
1838 TAATCATACCTCATTA 1859 TAATCATACCTCATTA 1186CATG 1960 1996 GATGATCACACATACA 1888 GACGAAGGCCTTTATG 1909 GACGAAGGCCTTTATG 1210 GGCGACTGTCAAGTAC 2010 2046 TGCGACGATTAAGTAC 1938 GGCGACTGTCAAGTAC 1959 GGCGACTGTCAAGTAC 1959 GGCGACTGTCAAGTAC 2050 ATAATGGTTTTGGTAA 2060 2096 ATGATGGATGGGGGCAA	TGACCATTAC TGACCATTAC AGGCACCTAA 1970 AAGCTCCAAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA AGGCACCTAA COCO TACGTAGAAC TATGTCGAAC TATGTCGAAC CAAATTGGGC CGCTAGCGAC CGCTAGCGAC CGCTAGCGAC CGCTAGCGAC	CATAACATC CATAACATC CATAACATT GGGGTATAC 1980 TGGCTATAC GGGGTATAC GGGGTATAC ATCCAGACG ATCCAAACG CATGTTTTX	A A A T T T G A G T A A A T T T G A G T 1990 C T T G G A A G A T T C T T G A G G A T T C T T G A G G A T A A C G T C C G C A 2040 A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A C G T C C G C A A A A A A C A A G A A G A A A C A A A A A A A A A A T A A G G	G G T T T phtB.seq G G T T T phtD.SEC phtE.SEC C T T T T Majority 2000 T T G T T phtB.seq C T T T T phtB.seq C T T T T phtD.SEO phtE.SEO T T C A G Majority 2050 T T C A G phtB.seq T T C A G phtB.seq T T C A G phtB.seq T T C A A phtE.SEO A A G A T Majority 2100 A A G A T Majority 2100 A A G A C phtB.seq T A G A C phtB.seq T A G A C phtB.seq

14/17

Figure 7(g)

C-A A G C C A G T A A A C C T J	ATGAAGATGAG	AAACATGACCA	AGTAAG-GAG Majority
2110	2120	2130	2140 2150
2143 CA C J	GTGAAGAT	C C A	AATAAG phtA.SEQ
2038 CAAGCTGATA CCJ			AGCGAG-GAGAA phtB.seq
2059 CAAGACAGTAAACCT	SATGAAGATAAG	GAACATGATGA	AGTAAGTGAGCC phtD.SEQ
1284 C A A T C C A G - G A A C T T C	: A C A T G A G .	A A C A T G A	phtE.SEQ
A - C T C A C - G A A	TGAAGAAGA	-	- TTTXXXTCCT - Majority
2160	2170	2180	2190 2200
2164			-TTCAAA pheA.SEQ
2079 ACCTCAGACAGAAAA			
2109 AACTCACCCTGAATCT			•
1314	XGXXGX	TGGATACGG	ATTTGA-TGCT- phtE.SEQ
		•	- A - A C A A Majority
- AGCAGATAAACCGT			
2210	2220	2230	2240 2250
2173 G C G G A T G A A	•		phtA.SEQ
2114 G A G A A G A G A A A C C G C A 2159 C A G C Á G A T A A T C T T T A			AAAACCAACAGA phtB.seq
1339			phtz.SEQ
G-AGCTGGAGGAAXCX	CCAGATGAGTC	AGAAGTXCCTC	AAGTAGAGACTG Majority
2260	2270	2280	2290 2300
2189 TAGAGGAAACA			
Troy Troy Country to	CCTGCTGAGCC	AGAAGTCCCTC	AAGTAGAGACTG phtA.SEQ
2163 GGAACCAGAAGAATCA	CCAGAGGAATC	AGAAGAACCTC	AGGTCGAGACTG phtB.seq
2163 GGAACCAGAAGAATCA 2209 GAAGCTGAAGATACCA	CCAGAGGAATC;	AGAAGAACCTC	AGGTCGAGACTG phtB.seq
2163 GGAACCAGAAGAATCA	CCAGAGGAATC	AGAAGAACCTC	AGGTCGAGACTG phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A	CCAGAGGAATC		AGGTCGAGACTG phtB.seq
2163 GGAACCAGAAGAATCA 2209 GAAGCTGAAGATACCA 1351 GCTGAAGA	CCAGAGGAATC	AGAAGAACCTC TGAAATTCCTC CXGAGGTTTTG	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310	CCAGAGGAATCA C-AGATGAGGC TGAATC ACTXAXAGAXG	AGAAGAACCTC FGAAATTCCTC CXGAGGTTTTG	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G T A G A A G C C C A	CCAGAGGAATCA C-AGATGAGGC TGAATCA ACTXAXAGAXG	AGAAGAACCTC TGAAATTCCTC CXGAGGTTTTG 2330 CAGAAGTTTTG	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A	CCAGAGGAATCA C-AGATGAGGC TGAATC ACTXAXAGAXG	A G A A G A A C C T C T G A A A T T C C T C C X G A G G T T T T G 2330 C A G A A G T T T T G C T G A A G A T T T A	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A	2320 ACTCAAAGAAG ACTCAAAGAAG ACTGAGAGAGG	A G A A G A A C C T C T G A A A T T C C T C C X G A G G T T T T G 2330 C A G A A G T T T T G C T G A A G A T T T A	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGGAAAAGTA phtA.SEQ CTTGGAAAAGTA phtB.seq CTAGAAAAATC phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG GATAGCAGATG	AGAAGAACCTC TGAAATTCCTC 2330 CAGAAGTTTTG CTGAAGATTTA CGGAAGGTTTTG	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A A G T T G A A G C C C A 2310 2234 A A A A A G T T G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG	2330 CAGAAGTTTTG CTGAAGTTTTG CTGAAGATTTA CGGAAGATTTA CGGAGGTTTTG	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAATC phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C C C A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG GATAGCAGATG AAXCCAATGCX 2370	AGAAGATTCCTC CXGAGGTTTTG 2330 CAGAAGTTTTG CTGAAGATTTA CGGAGGCCTTG ACGGAGACTCT	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq CTAGAAAAAAAATC phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A	CCAGAGGAATC C-AGATGAGGC TGAATC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX 2370	2330 CAGAAGTTTTG CTGAAGTTTTG CTGAAGATTTA CGGAAGATTTA CGGAGACTCT 2380	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq AACTGGTTTAAA Majority 2390 2400 AGCTGGTTTACG phtA.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A A G T A G A A G C C C A 2213 A A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A 2263 C A G G A T C C A A T T A T C A	2320 ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG AAXCCAATGCA AAXCCAATGCA	2330 CAGAAGTTTTG CTGAAGTTTTG CTGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAGTA phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAAATC phtB.seq CACAGGATTAAAA Majority 2390 2400 AGCTGGTTTAAA phtB.seq CACAGGATTAAAA phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G G T T G A A G C C C A 2213 A A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A 2263 C A G G A T C C T A G T A T T A T C A 2308 A C A G A T C C T A G T A T T A	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX AAXCCAATGCX AAGCCAATGCX AGTCCAATGCC	2330 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CAACTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA phtB.seq CACAGGATTAAA phtB.seq CACAGGATTAAA phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A A G T A G A A G C C C A 2213 A A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A 2263 C A G G A T C C A A T T A T C A	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX AAXCCAATGCX AAGCCAATGCX AGTCCAATGCC	2330 CAGAAGTTTTG CTGAAGTTTTG CTGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA CGGAAGATTTA	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CAACTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA phtB.seq CACAGGATTAAA phtB.seq CACAGGATTAAA phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A T A C C A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T A T A C 2360 2284 A C G G A T T C T A G T C T G A 2308 A C A G A T C C T A G T A T T A 1375 A T G A G T C	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX 2370 AAGCCAATGCA AGTCCAATGCA AGTCCAATGCA	2330 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CAACTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA phtB.seq CACAGGATTAAA phtB.seq CACAGGATTAAA phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A T A C C A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T A T A C 2360 2284 A C G G A T T C T A G T C T G A 2308 A C A G A T C C T A G T A T T A 1375 A T G A G T C	CCAGAGGAATC C-AGATGAGGC ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX 2370 AAGCCAATGCA AGTCCAATGCA AGTCCAATGCA	2330 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300 2300	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A 2308 A C A G A T C C T A G T A T T A 1375 A T G A G T C A A A T A A T T T X C T T C T T 2410	2320 ACTXAXAGAXG 2320 ACTXAXAGAXG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370	2330 2330 2330 2330 2330 2330 2330 2330 2330 2330 2330 2330 2430	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A X G T T G A A G C X A A 2310 2234 A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X A 2360 2284 A C G G A T T C T A G T C T G A 2308 A C A G A T C C T A G T A T T X 1375 A T G A G T C A A A T A A T T T X C T T C T T	2320 ACTXAXAGAXG 2320 ACTXAXAGAXG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG ACTGAGAGAGG AAXCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370	2330 2330	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ phtE.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAAATC phtB.seq CTAGAAAAAATC phtE.SEQ XACTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA Majority CACAGGATTAAA phtB.seq GACTGGTTTAAA phtB.seq GACTGGTCTAAA phtB.seq
2163 G G A A C C A G A A G A A T C A 2209 G A A G C T G A A G A T A C C A 1351 G C T G A A G A A A A A A G T T G A A G C C C A 2310 2234 A A A A A A G T A G A A G C C C A 2213 A A A A G G T T G A A G A A A A 2258 C T G T T A T T A A C G C T A A 1365 A C G G A T C C T A G T A T X X 2360 2284 A C G G A T T C T A G T C T G A 2263 C A G G A T C C A A T T A T C A 2308 A C A G A T C C T A G T A T T X 1375 A T G A G T C A A A T A A T T T X C T T C T T 2410 2334 A A A T A A T T T G A C T C T T	2320 ACTXAXAGAXG 2320 ACTCAAAGAAG ACTGAGAGAGG AAXCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX 2370 AAGCCAATGCX AGGCACCAAGGA CGACACCAAGGA	2330 2330	AGGTCGAGACTG phtB.seq AAGTAGAGAATT phtD.SEQ phtZ.SEQ CTTGXAAAAGTC Majority 2340 2350 CTTGCGAAAAGTA phtA.SEQ CTTGGAAAAATC phtB.seq CTAGAAAAATC phtB.seq CTAGAAAAATC phtZ.SEQ XACTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA Majority 2390 2400 AGCTGGTTTAAA phtB.seq GACAGGATTAAA phtB.seq GACAGGATTAAA phtB.seq GACTGGTCTAAA phtB.seq CACAGGAAAG AAG phtA.SEQ

15/17

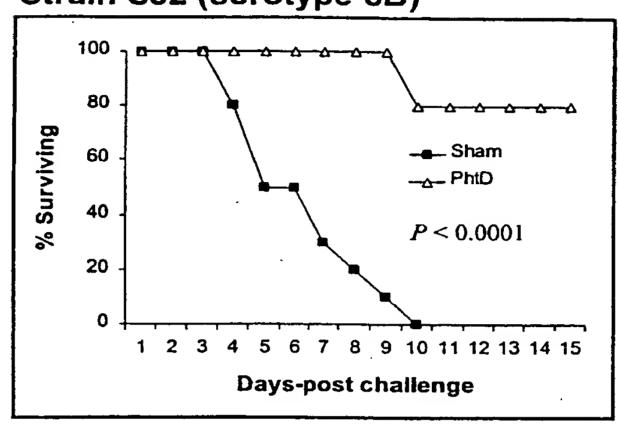
Figure 7(h)

	2	:)		G	λ	λ	λ	G	,	(_	Ţ	λ	T	. 7	•	3	G	¢	T	T	Ţ	G	T	7	,	. /		\ (3 (3)		: >	٥ ١	7	. ,		L X	7	_	λ	A	G	G	T			<u>-</u>	<u>.</u>	- (: 1	rī	Majori
											24		0									2	47	0								_ :	24	80					_			2	490)								25(00
2384	C	: A	. (G	λ	λ	λ	λ	. ,	. 1		T	λ	C	7		- 1	3	C	G	T	T	G	T	I	, y	. ,				; C	· A	. ,	G	7	,		T	C	-	_	-	_	-	_					- (7	<u> </u>	phtA.S
2363	C	7	. (G	X	λ	λ	λ		. (= '	T	λ	T	. 3	•	; (3	כ '	T	T	Ŧ	X	T	T	, y	J	J	, C	G	3 7	a	X	G	T	. 7	l A	G	T	A	A	A	G	3	T a	A (G J	N J	١ (3 C	: 7	T	phtB.s
2408	I		- '	_	λ	_		G	-		: '	Ť	C	T	7		: (3 (כ '	T	T	T	G	T	T	y				_	; y	•		G	•	-	-	-	-	-	-	-	-	-							. <u>-</u>	. c	phtD.S
1409	-	À	. (3	X	λ	G	G	y	. (•	T	-	-	7	. (3 2	R.	2	-	-	-	-	-	-	-	-		G	X	, y	G	λ	G	C		A	λ	T	T	λ	λ	G (3 :	r ·	- •					-	-	phtE.S
	_	_				_	_	_				_	_	_	_	_		_		_	_	_	_		_		_	_	_																			•.	۲,				
	<u> </u>	. A		_	_	G	<u></u>	-5		_		T	C	T	_	-		=	- (<u> </u>		Ŧ.	수	_	G	_	C			_	<u>. , y</u>	<u>, y</u>		<u>,</u>	_	7	<u>_</u>		T	λ	T	G	<u>G</u>	<u> </u>	A J	<u> </u>	<u> </u>	, c	<u>: c</u>	<u> </u>	<u> </u>	<u> </u>	_ Majori
							_			:	25	1	0									2	52	0							,	2	5	30								25	40)							;	255	50
2423	C	À		•	-	_	_	-	_	-	•	r	C	T	G	-		•		•	-	T	λ	λ	G	-	-	_	-	-	_	-	_	-	-	_	-	-	T	λ	X	G	G J	7		\		<u>_</u>	7				- phtA.SI
2413				3	3	C	C	G	X			r	Ť	T	Ø	C	. (2 2	•	2 (C	X	G	G	λ	C	À	À	C	. A	λ	T	A	C	T	A	T		A													phtB.se
2441	_			•	-	_	-	-	-	-	•	-	C	C	G	C	; (. (2	כ '	T	X	-	-	-	-	-	-	-	· -	-	-	-	-	-	-	-	T	λ	T	A I	3 7	r J									phtD.S
1440	-	-	-	•	- (G	С	G	-	-	•	-	-	-	-	-	•	• •	•	•	-	-	-	-	_	-	С	λ	λ	À		A	C	λ	T	T	-	-	-	-	T	A (G										phtz.si
•	<u> </u>	<u> </u>			<u> </u>	<u>c</u>	T	λ	x	X	_																																										Majori
2445	-	<u> </u>	_		_		T	λ	<u> </u>		-																																										obes ci
2463	λ	A	A	ı			Ţ	A	T	T	•															•	,																										phtA.SI
246B	λ	G	-			-	-	-	C	C	•																															•											phtD.SI
1455																																																					phtE.SI

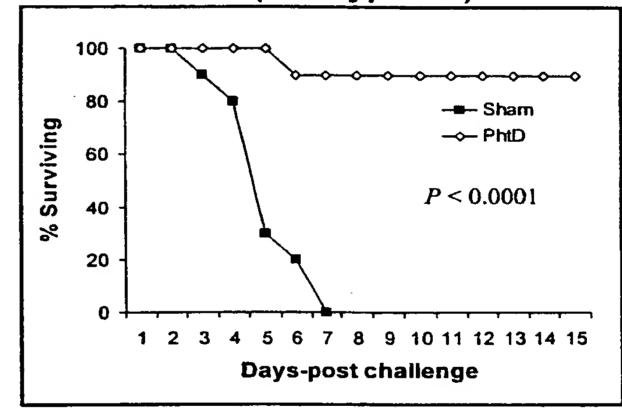
16/17

Figure 8

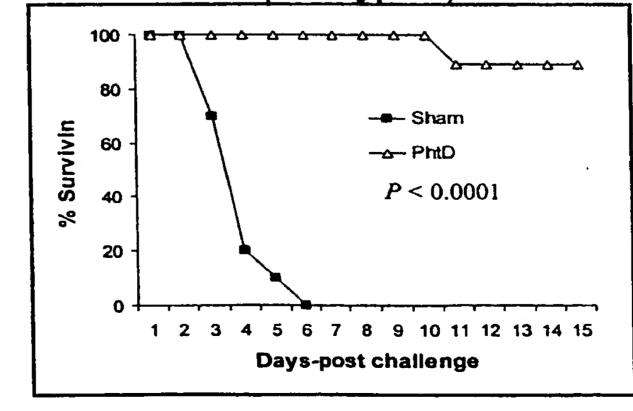
A. Strain SJ2 (serotype 6B)



B. Strain EF6796 (serotype 6A)



C. Strain EF5668 (serotype 4)



17/17

SEQUENCE LISTING

<110> Johnson, Leslie S. Koenig, Scott Adamou, John E. <120> Streptococcus Pneumoniae and Immunogenic Fragments for Vaccines <130> 469201-444 <140> <141> <150> 60/113,048 <151> 1998-12-21 <160> 11 <170> PatentIn Ver. 2.0 <210> 1 <211> 36 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Forward primer used in amplification of the Sp36 gene sequence. <400> 1 atcggatcct tcttacgagt tgggactgta tcaagc 36 <210> 2 <211> 35 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Forward primer used in amplification of the Sp36 gene sequence. <400> 2 atcggatcca ctgtatcaag ctagaacggt taagg 35 <210> 3 <211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Reverse primer used in amplification of the Sp36 gene sequence.

<400> 3

agtcaagctt gtttattttt tccttactta cagatgaagg

40

<210> 4

<211> 838

<212> PRT

<213> Streptococcus pneumoniae

<400> 4

Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala 1 5 10 15

Leu Ser Val Cys Ser Tyr Glu Leu Gly Arg His Gln Ala Gly Gln Val
20 25 30

Lys Lys Glu Ser Asn Arg Val Ser Tyr Ile Asp Gly Asp Gln Ala Gly
35 40 45

Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly
50 55 60

Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val
65 70 75 80

Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
85 90 95

Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
100 105 110

Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 115 120 125

Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140

Asp Asn Ile Arg Thr Lys Glu Glu Ile Lys Arg Gln Lys Gln Glu His
145 150 155 160

Ser His Asn His Gly Gly Gly Ser Asn Asp Gln Ala Val Val Ala Ala 165 170 175

Arg Ala Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu Ala Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu . Asn His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu Asn Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu Arg His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His Phe Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg Ile Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro Glu Gln Pro Ser Pro Gln Ser Thr Pro Glu Pro Ser Pro Ser Pro Gln Pro Ala Pro Asn Pro Gln Pro Ala Pro Ser Asn Pro Ile Asp Glu Lys Leu Val Lys Glu Ala Val Arg Lys Val Gly Asp Gly Tyr Val Phe Glu Glu Asn Gly Val Ser Arg Tyr Ile Pro Ala Lys Asp Leu Ser Ala

Glu Thr Ala Ala Gly Ile Asp Ser Lys Leu Ala Lys Gln Glu Ser Leu

Ser His Lys Leu Gly Ala Lys Lys Thr Asp Leu Pro Ser Ser Asp Arg
435 440 445

- Glu Phe Tyr Asn Lys Ala Tyr Asp Leu Leu Ala Arg Ile His Gln Asp 450 455 460
- Leu Leu Asp Asn Lys Gly Arg Gln Val Asp Phe Glu Ala Leu Asp Asn 465 470 480
- Leu Leu Glu Arg Leu Lys Asp Val Pro Ser Asp Lys Val Lys Leu Val
 485 490 495
- Asp Asp Ile Leu Ala Phe Leu Ala Pro Ile Arg His Pro Glu Arg Leu 500 510
- Gly Lys Pro Asn Ala Gln Ile Thr Tyr Thr Asp Asp Glu Ile Gln Val 515 520 525
- Ala Lys Leu Ala Gly Lys Tyr Thr Thr Glu Asp Gly Tyr Ile Phe Asp 530 540
- Pro Arg Asp Ile Thr Ser Asp Glu Gly Asp Ala Tyr Val Thr Pro His 545 550 550 560
- Met Thr His Ser His Trp Ile Lys Lys Asp Ser Leu Ser Glu Ala Glu
 565 570 575
- Arg Ala Ala Gln Ala Tyr Ala Lys Glu Lys Gly Leu Thr Pro Pro 580 585 590
- Ser Thr Asp His Gln Asp Ser Gly Asn Thr Glu Ala Lys Gly Ala Glu

 595
 600
 605
- Ala Ile Tyr Asn Arg Val Lys Ala Ala Lys Lys Val Pro Leu Asp Arg 610 620
- Met Pro Tyr Asn Leu Gln Tyr Thr Val Glu Val Lys Asn Gly Ser Leu 625 630 635 640
- Ile Ile Pro His Tyr Asp His Tyr His Asn Ile Lys Phe Glu Trp Phe 645 650 655
- Asp Glu Gly Leu Tyr Glu Ala Pro Lys Gly Tyr Thr Leu Glu Asp Leu 660 670
- Leu Ala Thr Val Lys Tyr Tyr Val Glu His Pro Asn Glu Arg Pro His 675 680 685

Ser Asp Asn Gly Phe Gly Asn Ala Ser Asp His Val Arg Lys Asn Lys 690 695 700

Val Asp Gln Asp Ser Lys Pro Asp Glu Asp Lys Glu His Asp Glu Val
705 710 715 720

Ser Glu Pro Thr His Pro Glu Ser Asp Glu Lys Glu Asn His Ala Gly
725 730 735

Leu Asn Pro Ser Ala Asp Asn Leu Tyr Lys Pro Ser Thr Asp Thr Glu
740 745 750

Glu Thr Glu Glu Glu Ala Glu Asp Thr Thr Asp Glu Ala Glu Ile Pro 755 760 765

Gln Val Glu Asn Ser Val Ile Asn Ala Lys Ile Ala Asp Ala Glu Ala 770 780

Leu Leu Glu Lys Val Thr Asp Pro Ser Ile Arg Gln Asn Ala Met Glu
785 790 795 800

Thr Leu Thr Gly Leu Lys Ser Ser Leu Leu Leu Gly Thr Lys Asp Asn 805 810 815

Asn Thr Ile Ser Ala Glu Val Asp Ser Leu Leu Ala Leu Leu Lys Glu 820 825 830

Ser Gln Pro Ala Pro Ile 835

<210> 5

<211> 2531

<212> DNA

<213> Streptococcus pneumoniae

<400> 5

atgaaaatta ataaaaata tctagcaggt tcagtggcag tccttgccct aagtgttgt 60 tcctatgaac ttggtcgtca ccaagctggt caggttaaga aagagtctaa tcgagtttct 120 tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180 aagagggagg ggatcaacgc cgaacaaatc gtcatcaaga ttacggatca aggttatgtg 240 acctctcatg gagaccatta tcattactat aatggcaagg tcccttatga tgccatcatc 300 agtgaagagc tcctcatgaa agatccgaat tatcagttga aggattcaga cattgtcaat 360 gaaatcaagg gtggttatgt tatcaaggta gatggaaaat actatgtta ccttaaggat 420 gcagctcatg cggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacac 480 agtcataaca cggatggtgg ttctaacgat caagcagtag ttgcagcag agcccaagga 540 cgctatacaa cggatgatgg ttatatcttc aatgcatca ttcctaagaa ttgcatcaga tgagttatca 660 gatgcttata tcgttcctca cggcgaccat taccattaca ttcctaagaa tgagttatca 660

```
gctagcgagt tagctgctgc agaagcctat tggaatggga agcagggatc tcgtccttct 720
tcaagttcta gttataatgc aaatccagct caaccaagat tgtcagagaa ccacaatctg 780
actgtcactc caacttatca tcaaaatcaa ggggaaaaca tttcaagcct tttacgtgaa 840
ttgtatgcta aacccttatc agaacgccat gtggaatctg atggccttat tttcgaccca 900
gegeaaatea caagtegaae egeeagaggt gtagetgtee eteatggtaa ceattaceae 960
tttatccctt atgaacaaat gtctgaattg gaaaaacgaa ttgctcgtat tattcccctt 1020
cgttatcgtt caaaccattg ggtaccagat tcaagaccag aacaaccaag tccacaatcg 1080
actccggaac ctagtccaag tccgcaacct gcaccaaatc ctcaaccagc tccaagcaat 1140
ccaattgatg agaaattggt caaagaagct gttcgaaaag taggcgatgg ttatgtcttt 1200
gaggagaatg gagtttctcg ttatatccca gccaaggatc tttcagcaga aacagcagca 1260
ggcattgata gcaaactggc caagcaggaa agtttatctc ataagctagg agctaagaaa 1320
actgacctcc catctagtga tcgagaattt tacaataagg cttatgactt actagcaaga 1380
attcaccaag atttacttga taataaaggt cgacaagttg attttgaggc tttggataac 1440
ctgttggaac gactcaagga tgtcccaagt gataaagtca agttagtgga tgatattctt 1500
geettettag eteegatteg teatecagaa egtttaggaa aaccaaatge geaaattace 1560
tacactgatg atgagattca agtagccaag ttggcaggca agtacacaac agaagacggt 1620
tatatctttg atcctcgtga tataaccagt gatgaggggg atgcctatgt aactccacat 1680
atgacccata gccactggat taaaaaagat agtttgtctg aagctgagag agcggcagcc 1740
caggettatg ctaaagagaa aggtttgace eeteettega cagaceatea ggatteagga 1800
aatactgagg caaaaggagc agaagctatc tacaaccgcg tgaaagcagc taagaaggtg 1860
ccacttgatc gtatgcctta caatcttcaa tatactgtag aagtcaaaaa cggtagttta 1920
atcatacctc attatgacca ttaccataac atcaaatttg agtggtttga cgaaggcctt 1980
tatgaggcac ctaaggggta tactcttgag gatcttttgg cgactgtcaa gtactatgtc 2040
gaacatccaa acgaacgtcc gcattcagat aatggttttg gtaacgctag cgaccatgtt 2100
cgtaaaaata aggtagacca agacagtaaa cctgatgaag ataaggaaca tgatgaagta 2160
agtgagccaa ctcaccctga atctgatgaa aaagagaatc acgctggttt aaatccttca 2220
gcagataatc tttataaacc aagcactgat acggaagaga cagaggaaga agctgaagat 2280
accacagatg aggctgaaat tcctcaagta gagaattctg ttattaacgc taagatagca 2340
gatgcggagg ccttgctaga aaaagtaaca gatcctagta ttagacaaaa tgctatggag 2400
acattgactg gtctaaaaag tagtcttctt ctcggaacga aagataataa cactatttca 2460
gcagaagtag atagtctctt ggctttgtta aaagaaagtc aaccggctcc tatatagtaa 2520
aagcttaagc c
                                                                  2531
```

<210> 6

<211> 484

<212> PRT

<213> Streptococcus pneumoniae

<400> 6

Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val

1 10 15

Ser Leu Ser Leu Cys Ala Tyr Ala Leu Asn Gln His Arg Ser Gln Glu 20 25 30

Asn Lys Asp Asn Asn Arg Val Ser Tyr Val Asp Gly Ser Gln Ser Ser

40
45

Gln Lys Ser Glu Asn Leu Thr Pro Asp Gln Val Ser Gln Lys Glu Gly
50 60

- Ile Gln Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val
 65 70 75 80
- Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
 85 90 95
- Asp Ala Leu Phe Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
 100 105 110
- Leu Lys Asp Ala Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Ile Ile 115 120 125
- Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140
- Asp Asn Val Arg Thr Lys Asp Glu Ile Asn Arg Gln Lys Gln Glu His 145 150 155 160
- Val Lys Asp Asn Glu Lys Val Asn Ser Asn Val Ala Val Ala Arg Ser 165 170 175
- Gln Gly Arg Tyr Thr Thr Asn Asp Gly Tyr Val Phe Asn Pro Ala Asp 180 185 190
- Ile Ile Glu Asp Thr Gly Asn Ala Tyr Ile Val Pro His Gly Gly His 195 200 205
- Tyr His Tyr Ile Pro Lys Ser Asp Leu Ser Ala Ser Glu Leu Ala Ala 210 215 220
- Ala Lys Ala His Leu Ala Gly Lys Asn Met Gln Pro Ser Gln Leu Ser 230 235 240
- Tyr Ser Ser Thr Ala Ser Asp Asn Asn Thr Gln Ser Val Ala Lys Gly
 245 250 255
- Ser Thr Ser Lys Pro Ala Asn Lys Ser Glu Asn Leu Gln Ser Leu Leu 260 265 270
- Lys Glu Leu Tyr Asp Ser Pro Ser Ala Gln Arg Tyr Ser Glu Ser Asp 275 280 285
- Gly Leu Val Phe Asp Pro Ala Lys Ile Ile Ser Arg Thr Pro Asn Gly
 290 295 300

Val Ala Ile Pro His Gly Asp His Tyr His Phe Ile Pro Tyr Ser Lys 305 310 315 320

Leu Ser Ala Leu Glu Glu Lys Ile Ala Arg Met Val Pro Ile Ser Gly
325 330 335

Thr Gly Ser Thr Val Ser Thr Asn Ala Lys Pro Asn Glu Val Val Ser 340 345 350

Ser Leu Gly Ser Leu Ser Ser Asn Pro Ser Ser Leu Thr Thr Ser Lys 355 360 365

Glu Leu Ser Ser Ala Ser Asp Gly Tyr Ile Phe Asn Pro Lys Asp Ile 370 380

Val Glu Glu Thr Ala Thr Ala Tyr Ile Val Arg His Gly Asp His Phe 385 390 395 400

His Tyr Ile Pro Lys Ser Asn Gln Ile Gly Gln Pro Thr Leu Pro Asn 405 410 415

Asn Ser Leu Ala Thr Pro Ser Pro Ser Leu Pro Ile Asn Pro Gly Thr
420 425 430

Ser His Glu Lys His Glu Glu Asp Gly Tyr Gly Phe Asp Ala Asn Arg 435 440 445

Ile Ile Ala Glu Asp Glu Ser Gly Phe Val Met Ser His Gly Asp His
450 455 460

Asn His Tyr Phe Phe Lys Lys Asp Leu Thr Glu Glu Gln Ile Lys Val 465 470 475 480

Arg Lys Asn Ile

<210> 7

<211> 1455

<212> DNA

<213> Streptococcus pneumoniae

<400> 7

atgaaattta gtaaaaata tatagcagct ggatcagctg ttatcgtatc cttgagtcta 60 tgtgcctatg cactaaacca gcatcgttcg caggaaaata aggacaataa tcgtgtctct 120 tatgtggatg gcagccagtc aagtcagaaa agtgaaaact tgacaccaga ccaggttagc 180 cagaaagaag gaattcaggc tgagcaaatt gtaatcaaaa ttacagatca gggctatgta 240 acgtcacacg gtgaccacta tcattactat aatgggaaag ttccttatga tgccctcttt 300

```
agtgaagaac tcttgatgaa ggatccaaac tatcaactta aagacgctga tattgtcaat 360
gaagtcaagg gtggttatat catcaaggtc gatggaaaat attatgtcta cctgaaagat 420
gcagctcatg ctgataatgt tcgaactaaa gatgaaatca atcgtcaaaa acaagaacat 480
gtcaaagata atgagaaggt taactctaat gttgctgtag caaggtctca gggacgatat 540
acgacaaatg atggttatgt ctttaatcca gctgatatta tcgaagatac gggtaatgct 600
tatatcgttc ctcatggagg tcactatcac tacattccca aaagcgattt atctgctagt 660
gaattagcag cagctaaagc acatctggct ggaaaaaata tgcaaccgag tcagttaagc 720
tattetteaa cagetagtga caataacaeg caatetgtag caaaaggate aactagcaag 780
ccagcaaata aatctgaaaa tctccagagt cttttgaagg aactctatga ttcacctagc 840
gcccaacgtt acagtgaatc agatggcctg gtctttgacc ctgctaagat tatcagtcgt 900
acaccaaatg gagttgcgat tccgcatggc gaccattacc actttattcc ttacagcaag 960
ctttctgcct tagaagaaaa gattgccaga atggtgccta tcagtggaac tggttctaca 1020
gtttctacaa atgcaaaacc taatgaagta gtgtctagtc taggcagtct ttcaagcaat 1080
ccttcttctt taacgacaag taaggagctc tcttcagcat ctgatggtta tatttttaat 1140
ccaaaagata tcgttgaaga aacggctaca gcttatattg taagacatgg tgatcatttc 1200
cattacattc caaaatcaaa tcaaattggg caaccgactc ttccaaacaa tagtctagca 1260
acaccttctc catctctcc aatcaatcca ggaacttcac atgagaaaca tgaagaagat 1320
ggatacggat ttgatgctaa tcgtattatc gctgaagatg aatcaggttt tgtcatgagt 1380
cacggagacc acaatcatta tttcttcaag aaggacttga cagaagagca aattaaggtg 1440
cgcaaaaaca tttag
                                                                  1455
```

<210> 8

<211> 819

<212> PRT

<213> Streptococcus pneumoniae

<400> 8

Met Lys Ile Asn Lys Lys Tyr Leu Val Gly Ser Ala Ala Ala Leu Ile 1 5 10 15

Leu Ser Val Cys Ser Tyr Glu Leu Gly Leu Tyr Gln Ala Arg Thr Val
20 25 30

Lys Glu Asn Asn Arg Val Ser Tyr Ile Asp Gly Lys Gln Ala Thr Gln
35 40 45

Lys Thr Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly Ile
50 60

Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val Thr
65 70 75 80

Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr Asp 85 90 95

Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Lys Leu 100 105 110

Lys Asp Glu Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Val Ile Lys
115 120 125

- Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala Asp 130 135 140
- Asn Val Arg Thr Lys Glu Glu Ile Asn Arg Gln Lys Gln Glu His Ser 145 150 155 160
- Gln His Arg Glu Gly Gly Thr Pro Arg Asn Asp Gly Ala Val Ala Leu 165 170 175
- Ala Arg Ser Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn 180 185 190
- Ala Ser Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His 195 200 205
- Gly Asp His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu 210 215 220
- Leu Ala Ala Glu Ala Phe Leu Ser Gly Arg Gly Asn Leu Ser Asn 225 230 235 240
- Ser Arg Thr Tyr Arg Arg Gln Asn Ser Asp Asn Thr Ser Arg Thr Asn 245 250 255
- Trp Val Pro Ser Val Ser Asn Pro Gly Thr Thr Asn Thr Asn Thr Ser 260 265 270
- Asn Asn Ser Asn Thr Asn Ser Gln Ala Ser Gln Ser Asn Asp Ile Asp 275 280 285
- Ser Leu Leu Lys Gln Leu Tyr Lys Leu Pro Leu Ser Gln Arg His Val 290 295 300
- Glu Ser Asp Gly Leu Val Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr 305 310 315
- Ala Arg Gly Val Ala Val Pro His Gly Asp His Tyr His Phe Ile Pro 325 330 335
- Tyr Ser Gln Met Ser Glu Leu Glu Glu Arg Ile Ala Arg Ile Ile Pro 340 . 350
- Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro Glu Gln
 355 360 365

Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Gly Pro Gln Pro Ala 370 375 380

- Pro Asn Leu Lys Ile Asp Ser Asn Ser Ser Leu Val Ser Gln Leu Val
 385 390 395 400
- Arg Lys Val Gly Glu Gly Tyr Val Phe Glu Glu Lys Gly Ile Ser Arg
 405
 410
 415
- Tyr Val Phe Ala Lys Asp Leu Pro Ser Glu Thr Val Lys Asn Leu Glu
 420 425 430
- Ser Lys Leu Ser Lys Gln Glu Ser Val Ser His Thr Leu Thr Ala Lys
 435
 440
 445
- Lys Glu Asn Val Ala Pro Arg Asp Gln Glu Phe Tyr Asp Lys Ala Tyr
 450 455 460
- Asn Leu Leu Thr Glu Ala His Lys Ala Leu Phe Glu Asn Lys Gly Arg
 465 470 475 480
- Asn Ser Asp Phe Gln Ala Leu Asp Lys Leu Leu Glu Arg Leu Asn Asp
 485
 490
 495
- Glu Ser Thr Asn Lys Glu Lys Leu Val Asp Asp Leu Leu Ala Phe Leu
 500 505 510
- Ala Pro Ile Thr His Pro Glu Arg Leu Gly Lys Pro Asn Ser Gln Ile 515 520 525
- Glu Tyr Thr Glu Asp Glu Val Arg Ile Ala Gln Leu Ala Asp Lys Tyr
 530 540
- Thr Thr Ser Asp Gly Tyr Ile Phe Asp Glu His Asp Ile Ile Ser Asp 545 550 550
- Glu Gly Asp Ala Tyr Val Thr Pro His Met Gly His Ser His Trp Ile
 565 570 575
- Gly Lys Asp Ser Leu Ser Asp Lys Glu Lys Val Ala Ala Gln Ala Tyr 580 585 590
- Thr Lys Glu Lys Gly Ile Leu Pro Pro Ser Pro Asp Ala Asp Val Lys
 595 600 605
- Ala Asn Pro Thr Gly Asp Ser Ala Ala Ala Ile Tyr Asn Arg Val Lys
 610 620

Gly Glu Lys Arg Ile Pro Leu Val Arg Leu Pro Tyr Met Val Glu His 625 630 635 640

Thr Val Glu Val Lys Asn Gly Asn Leu Ile Ile Pro His Lys Asp His 655

Tyr His Asn Ile Lys Phe Ala Trp Phe Asp Asp His Thr Tyr Lys Ala 660 670

Pro Asn Gly Tyr Thr Leu Glu Asp Leu Phe Ala Thr Ile Lys Tyr Tyr 675 680 685

Val Glu His Pro Asp Glu Arg Pro His Ser Asn Asp Gly Trp Gly Asn 690 695 700

Ala Ser Glu His Val Leu Gly Lys Lys Asp His Ser Glu Asp Pro Asn 705 710 715 720

Lys Asn Phe Lys Ala Asp Glu Glu Pro Val Glu Glu Thr Pro Ala Glu
725 730 735

Pro Glu Val Pro Gln Val Glu Thr Glu Lys Val Glu Ala Gln Leu Lys
740 745 750

Glu Ala Glu Val Leu Leu Ala Lys Val Thr Asp Ser Ser Leu Lys Ala 755 760 765

Asn Ala Thr Glu Thr Leu Ala Gly Leu Arg Asn Asn Leu Thr Leu Gln
770 780

Ile Met Asp Asn Asn Ser Ile Met Ala Glu Ala Glu Lys Leu Leu Ala 785 790 795 800

Leu Leu Lys Gly Ser Asn Pro Ser Ser Val Ser Lys Glu Lys Ile Asn 805 810 815

Lys Leu Asn

<210> 9

<211> 2451

<212> DNA

<213> Streptococcus pneumoniae

<400> 9

atgaaaatta ataagaaata ccttgttggt tctgcggcag ctttgatttt aagtgtttgt 60 tcttacgagt tgggactgta tcaagctaga acggttaagg aaaataatcg tgtttcctat 120

PCT/US99/30390 WO 00/37105

```
atagatggaa aacaagcgac gcaaaaaacg gagaatttga ctcctgatga ggttagcaag 180
cgtgaaggàa tcaatgctga gcaaatcgtc atcaagataa cagaccaagg ctatgtcact 240
tcacatggcg accactatca ttattacaat ggtaaggttc cttatgacgc tatcatcagt 300
gaagaattac tcatgaaaga tccaaactat aagctaaaag atgaggatat tgttaatgag 360
gtcaagggtg gatatgttat caaggtagat ggaaaatact atgtttacct taaggatgct 420
gcccacgcgg ataacgtccg tacaaaagag gaaatcaatc gacaaaaaca agagcatagt 480
caacatcgtg aaggtggaac tccaagaaac gatggtgctg ttgccttggc acgttcgcaa 540
ggacgctata ctacagatga tggttatatc tttaatgctt ctgatatcat agaggatact 600
ggtgatgctt atatcgttcc tcatggagat cattaccatt acattcctaa gaatgagtta 660
tcagctagcg agttggctgc tgcagaagcc ttcctatctg gtcgaggaaa tctgtcaaat 720
tcaagaacct atcgccgaca aaatagcgat aacacttcaa gaacaaactg ggtaccttct 780
gtaagcaatc caggaactac aaatactaac acaagcaaca acagcaacac taacagtcaa 840
gcaagtcaaa gtaatgacat tgatagtctc ttgaaacagc tctacaaact gcctttgagt 900
caacgacatg tagaatctga tggccttgtc tttgatccag cacaaatcac aagtcgaaca 960
gctagaggtg ttgcagtgcc acacggagat cattaccact tcatccctta ctctcaaatg 1020
tctgaattgg aagaacgaat cgctcgtatt attccccttc gttatcgttc aaaccattgg 1080
gtaccagatt caaggccaga acaaccaagt ccacaaccga ctccggaacc tagtccaggc 1140
ccgcaacctg caccaaatct taaaatagac tcaaattctt ctttggttag tcagctggta 1200
cgaaaagttg gggaaggata tgtattcgaa gaaaagggca tctctcgtta tgtctttgcg 1260
aaagatttac catctgaaac tgttaaaaat cttgaaagca agttatcaaa acaagagagt 1320
gtttcacaca ctttaactgc taaaaaagaa aatgttgctc ctcgtgacca agaattttat 1380
gataaagcat ataatctgtt aactgaggct cataaagcct tgtttgnaaa taagggtcgt 1440
aattctgatt tccaagcctt agacaaatta ttagaacgct tgaatgatga atcgactaat 1500
aaagaaaaat tggtagatga tttattggca ttcctagcac caattaccca tccagagcga 1560
cttggcaaac caaattctca aattgagtat actgaagacg aagttcgtat tgctcaatta 1620
gctgataagt atacaacgtc agatggttac atttttgatg aacatgatat aatcagtgat 1680
gaaggagatg catatgtaac gcctcatatg ggccatagtc actggattgg aaaagatagc 1740
ctttctgata aggaaaaagt tgcagctcaa gcctatacta aagaaaaagg tatcctacct 1800
ccatctccag acgcagatgt taaagcaaat ccaactggag atagtgcagc agctatttac 1860
aatcgtgtga aaggggaaaa acgaattcca ctcgttcgac ttccatatat ggttgagcat 1920
acagttgagg ttaaaaacgg taatttgatt attcctcata aggatcatta ccataatatt 1980
aaatttgctt ggtttgatga tcacacatac aaagctccaa atggctatac cttggaagat 2040
ttgtttgcga cgattaagta ctacgtagaa caccctgacg aacgtccaca ttctaatgat 2100
ggatggggca atgccagtga gcatgtgtta ggcaagaaag accacagtga agatccaaat 2160
aagaacttca aagcggatga agagccagta gaggaaacac ctgctgagcc agaagtccct 2220
caagtagaga ctgaaaaagt agaagcccaa ctcaaagaag cagaagtttt gcttgcgaaa 2280
gtaacggatt ctagtctgaa agccaatgca acagaaactc tagctggttt acgaaataat 2340
ttgactcttc aaattatgga taacaatagt atcatggcag aagcagaaaa attacttgcg 2400
ttgttaaaag gaagtaatcc ttcatctgta agtaaggaaa aaataaacta a
                                                                  2451
<210> 10
<211> 819
```

<212> PRT

<213> Streptococcus pneumoniae

<400> 10

Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Val Leu Ala 1 10 15

Leu Ser Val Cys Ser Tyr Glu Leu Gly Arg Tyr Gln Ala Gly Gln Asp
20 25 30

- Lys Lys Glu Ser Asn Arg Val Ala Tyr Ile Asp Gly Asp Gln Ala Gly
 35 40 45
- Gln Lys Ala Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly
 50 55 60
- Ile Asn Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val
 65 70 75 80
- Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
 85 90 95
- Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
 100 105 110
- Leu Lys Asp Ser Asp Ile Val Asn Glu Ile Lys Gly Gly Tyr Val Ile 115 120 125
- Lys Val Asn Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala 130 135 140
- Asp Asn Ile Arg Thr Lys Glu Glu Ile Lys Arg Gln Lys Gln Glu Arg 145 150 155 160
- Ser His Asn His Asn Ser Arg Ala Asp Asn Ala Val Ala Ala Arg 165 170 175
- Ala Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Asn Ala Ser 180 185 190
- Asp Ile Ile Glu Asp Thr Gly Asp Ala Tyr Ile Val Pro His Gly Asp 195 200 205
- His Tyr His Tyr Ile Pro Lys Asn Glu Leu Ser Ala Ser Glu Leu Ala 210 215 220
- Ala Ala Glu Ala Tyr Trp Asn Gly Lys Gln Gly Ser Arg Pro Ser Ser 225 230 235 240
- Ser Ser Ser Tyr Asn Ala Asn Pro Ala Gln Pro Arg Leu Ser Glu Asn 245 250 255
- His Asn Leu Thr Val Thr Pro Thr Tyr His Gln Asn Gln Gly Glu Asn 260 265 270

Ile Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu Arg 275 280 285

- His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr Ser 290 295 300
- Arg Thr Ala Arg Gly Val Ala Val Pro His Gly Asn His Tyr His Phe 305 310 315 320
- Ile Pro Tyr Glu Gln Met Ser Glu Leu Glu Lys Arg Ile Ala Arg Ile 325 330 335
- Ile Pro Leu Arg Tyr Arg Ser Asn His Trp Val Pro Asp Ser Arg Pro 340 350
- Glu Glu Pro Ser Pro Gln Pro Thr Pro Glu Pro Ser Pro Ser Pro Gln 355 360 365
- Pro Ala Pro Ser Asn Pro Ile Asp Gly Lys Leu Val Lys Glu Ala Val 370 375 380
- Arg Lys Val Gly Asp Gly Tyr Val Phe Glu Glu Asn Gly Val Ser Arg 385 390 395 400
- Tyr Ile Pro Ala Lys Asp Leu Ser Ala Glu Thr Ala Ala Gly Ile Asp
 405
 410
 415
- Ser Lys Leu Ala Lys Gln Glu Ser Leu Ser His Lys Leu Gly Thr Lys 420 425 430
- Lys Thr Asp Leu Pro Ser Ser Asp Arg Glu Phe Tyr Asn Lys Ala Tyr
 435
 440
 445
- Asp Leu Leu Ala Arg Ile His Gln Asp Leu Leu Asp Asn Lys Gly Arg 450 455 460
- Gln Val Asp Phe Glu Ala Leu Asp Asn Leu Leu Glu Arg Leu Lys Asp 465 470 475 480
- Val Ser Ser Asp Lys Val Lys Leu Val Glu Asp Ile Leu Ala Phe Leu
 485 490 495
- Ala Pro Ile Arg His Pro Glu Arg Leu Gly Lys Pro Asn Ala Gln Ile 500 505 510
- Thr Tyr Thr Asp Asp Glu Ile Gln Val Ala Lys Leu Ala Gly Lys Tyr
 515 520 525

Thr	Ala	Glu	Asp	Gly	Tyr	Ile	Phe	Asp	Pro	Arg	Asp	Ile	Thr	Ser	Asp
	530					535					540				

- Glu Gly Asp Ala Tyr Val Thr Pro His Met Thr His Ser His Trp Ile
 545 550 555 560
- Lys Lys Asp Ser Leu Ser Glu Ala Glu Arg Ala Ala Ala Gln Ala Tyr 565 570 575
- Ala Glu Glu Lys Gly Leu Thr Pro Pro Ser Thr Asp His Gln Asp Ser

 580 585 590
- Gly Asn Thr Glu Ala Lys Gly Ala Glu Ala Ile Tyr Asn Arg Val Lys
 595 600 605
- Ala Ala Lys Lys Val Pro Leu Asp Arg Met Pro Tyr Asn Leu Gln Tyr 610 620
- Thr Val Glu Val Lys Asn Gly Ser Leu Ile Ile Pro His Tyr Asp His 625 630 635
- Tyr His Asn Ile Lys Phe Glu Trp Phe Asp Glu Gly Leu Tyr Glu Ala 645 650 655
- Pro Lys Gly Tyr Thr Leu Glu Asp Leu Leu Ala Thr Val Lys Tyr Tyr 660 670
- Val Glu His Pro Asn Glu Arg Pro His Ser Asp Asn Gly Phe Gly Asn 675 680 685
- Ala Ser Asp His Val Gln Arg Asn Lys Asn Gly Gln Ala Asp Thr Asn 690 695 700
- Gln Thr Glu Lys Pro Ser Glu Glu Lys Pro Gln Thr Glu Lys Pro Glu 705 710 715 720
- Glu Glu Thr Pro Arg Glu Glu Lys Pro Gln Ser Glu Lys Pro Glu Ser
 725 730 735
- Pro Lys Pro Thr Glu Glu Pro Glu Glu Ser Pro Glu Glu Ser Glu Glu 740 745 750
- Pro Gln Val Glu Thr Glu Lys Val Glu Glu Lys Leu Arg Glu Ala Glu
 755 760 765
- Asp Leu Cly Lys Ile Gln Asp Pro Ile Ile Lys Ser Asn Ala Lys
 770 780

Glu Thr Leu Thr Gly Leu Lys Asn Asn Leu Leu Phe Gly Thr Gln Asp
785 790 795 800

Asn Asn Thr Ile Met Ala Glu Ala Glu Lys Leu Leu Ala Leu Leu Lys 805 810 815

Glu Ser Lys

<210> 11

<211> 2531

<212> DNA

<213> Streptococcus pneumoniae

<400> 11

atgaaaatta ataaaaata tctagcaggt tcagtggcag tccttgccct aagtgtttgt 60 tcctatgagc ttggacgtta ccaagctggt caggataaga aagagtctaa tcgagttgct 120 tatatagatg gtgatcaggc tggtcaaaag gcagaaaact tgacaccaga tgaagtcagt 180 aagagggagg ggatcaacgc cgaacaatt gttatcaaga ttacggatca aggttatgtg 240 acctctcatg gagaccatta tcattactat aatggcaagg ttccttatga tgccatcatc 300 agtgaagage teeteatgaa agateegaat tateagttga aggatteaga eattgteaat 360 gaaatcaagg gtggttatgt cattaaggta aacggtaaat actatgttta ccttaaggat 420 gcrgctcatg cggataatat tcggacaaaa gaagagatta aacgtcagaa gcaggaacgc 480 agtcataatc ataactcaag agcagataat gctgttgctg cagccagagc ccaaggacgt 540 tatacaacgg atgatgggta tatcttcaat gcatctgata tcattgagga cacgggtgat 600 gcttatatcg ttcctcacgg cgaccattac cattacattc ctaagaatga gttatcagct 660 agegagttag etgetgeaga ageetattgg aatgggaage agggateteg teettettea 720 agttctagtt ataatgcaaa tccagctcaa ccaagattgt cagagaacca caatctgact 780 gtcactccaa cttatcatca aaatcaaggg gaaaacattt caagcctttt acgtgaattg 840 tatgctaaac ccttatcaga acgccatgtg gaatctgatg gccttatttt cgacccagcg 900 caaatcacaa gtcgaaccgc cagaggtgta gctgtccctc atggtaacca ttaccacttt 960 atcccttatg aacaaatgtc tgaattggaa aaacgaattg ctcgtattat tccccttcgt 1020 tatcgttcaa accattgggt accagattca agaccagaag aaccaagtcc acaaccgact 1080 ccagaaccta gtccaagtcc gcaaccagct ccaagcaatc caattgatgg gaaattggtc 1140 aaagaagctg ttcgaaaagt aggcgatggt tatgtctttg aggagaatgg agtttctcgt 1200 tatatcccag ccaaggatct ttcagcagaa acagcagcag gcattgatag caaactggcc 1260 aagcaggaaa gtttatctca taagctagga actaagaaaa ctgacctccc atctagtgat 1320 cgagaatttt acaataaggc ttatgactta ctagcaagaa ttcaccaaga tttacttgat 1380 aataaaggtc gacaagttga ttttgaggct ttggataacc tgttggaacg actcaaggat 1440 gtctcaagtg ataaagtcaa gttagtggaa gatattcttg ccttcttagc tccgattcgt 1500 catccagaac gtttaggaaa accaaatgcg caaattacct acactgatga tgagattcaa 1560 gtagccaagt tggcaggcaa gtacacagca gaagacggtt atatctttga tcctcgtgat 1620 ataaccagtg atgaggggga tgcctatgta actccacata tgacccatag ccactggatt 1680 aaaaaagata gtttgtctga agctgagaga gcggcagccc aggcttatgc traagagaaa 1740 ggtttgaccc ctccttcgac agaccatcag gattcaggaa atactgaggc aaaaggagca 1800 gaagctatct acaaccgmgt gaaagcagct aagaaggtgc cacttgatcg tatgccttac 1860

aatcttcaat	atactgtaga	agtcaaaaac	ggtagtttaa	tcatacctca	ttatgaccat	1920
taccataaca	tcaaatttga	gtggtttgac	gaaggccttt	atgaggcacc	taaggggtat	1980
actcttgagg	atcttttggc	gactqtcaag	tactatgtcg	aacatccaaa	cgaacgtccg	2040
cattcagata	atggttttgg	taacgctagc	gaccatgttc	aaagaaacaa	aaatggtcaa	2100
gctgatacca	atcaaacgga	aaaaccaagc	gaggagaaac	ctcagacaga	aaaacctgag	2160
gaagaaaccc	ctcgagaaga	gaaaccgcaa	agcgagaaac	cagagtctcc	aaaaccaaca	2220
gaggaaccag	aagaatcacc	agaggaatca	gaagaacctc	aggtcgagac	tgaaaaggtt	2280
gaagaaaaac	tgagagaggc	tgaagattta	cttggaaaaa	tccaggatcc	aattatcaag	2340
tccaatgcca	aagagactct	cacaggatta	aaaaataatt	tactatttgg	cacccaggac	2400
aacaatacta	ttatggcaga	agctgaaaaa	ctattggctt	tattaaagga	gagtaagtaa	2460
aggtagaagc	ttaagggcga	atttggcacc	caggacaaca	atactattat	ggcagaagct	2520
gaaaaactat	t					2531

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7: WO 00/37105 (11) International Publication Number: **A3** A61K 39/09, 39/40, A61P 31/04 // C07K (43) International Publication Date: 14/315 29 June 2000 (29.06.00) (21) International Application Number: (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, PCT/US99/30390 BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, 21 December 1999 (21.12.99) GB, GD, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, (22) International Filing Date: KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, (30) Priority Data: SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN. 21 December 1998 (21.12.98) YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, 60/113,048 US SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, (71) Applicant: MEDIMMUNE, INC. [US/US]; 35 West Watkins Mill Road, Gaithersburg, MD 20878 (US). SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). (72) Inventors: JOHNSON, Leslie, S.; 13545 Ambassador Drive, Germantown, MD 20874 (US). KOENIG, Scott; 10732 Ralston Road, Rockville, MD 20852 (US). ADAMOU, **Published** John, E.; 20822 Shamrock Glen Circle, Germantown, MD With international search report. 20874 (US). (88) Date of publication of the international search report: (74) Agents: GRANT, Alan, J. et al.; Carella, Byrne, Bain, Gilfillan, 9 November 2000 (09.11.00) Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

(54) Title: STREPTOCOCCUS PNEUMONIAE PROTEINS AND IMMUNOGENIC FRAGMENTS FOR VACCINES

(57) Abstract

A vaccine composition is disclosed that comprises polypeptides and fragments of polypeptides containing histidine triad residues or coiled-coil regions, some of which polypeptides or fragments lie between 80 and 680 residues in length. Also disclosed are processes for preventing infection caused by S. pneumoniae comprising administering of vaccine compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand	·	
CM	Cameroon		Republic of Korea	PL	Poland .		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	Li	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

tional Application No

PCT/US 99/30390 A. CLASSIFICATION OF SUBJECT MATTER A61K39/40 //C07K14/315 A61K39/09 A61P31/04 IPC 7 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) STRAND, EPO-Internal C. DOCUMENTS CONSIDERED TO BE RELEVANT Category ° Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. X WO 98 18930 A (HUMAN GENOME SCIENCES INC 1-3 ;CHOI GIL H (US); HROMOCKYJ ALEX (US); J) 7 May 1998 (1998-05-07) page 59 -page 60 page 4, line 14 -page 5, line 16 P,X WO 99 15675 A (GREEN BRUCE A ; CHENG QI 1 - 3(US); FINKEL DAVID J (US); MASI AMY W (US)) 1 April 1999 (1999-04-01) claims 50-52,55-59 page 18, line 21 -page 19, line 12 Further documents are listed in the continuation of box C. Patent family members are listed in annex. Special categories of cited documents : "T" later document published after the international filing date or priority date and not in conflict with the application but "A" document defining the general state of the art which is not cited to understand the principle or theory underlying the considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention filing date cannot be considered novel or cannot be considered to "L" document which may throw doubts on priority claim(s) or involve an inventive step when the document is taken alone which is cited to establish the publication date of another "Y" document of particular relevance; the claimed invention citation or other special reason (as specified) cannot be considered to involve an inventive step when the "O" document referring to an oral disclosure, use, exhibition or document is combined with one or more other such docuother means ments, such combination being obvious to a person skilled in the art. document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 10 August 2000 17/08/2000 Name and mailing address of the ISA Authorized officer European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Covone, M

Form PCT/ISA/210 (second sheet) (July 1992)

Fax: (+31-70) 340-3016

6

INTERNATIONAL SEARCH REPORT

In: Itional Application No PCT/US 99/30390

C (Continua	730390				
Category Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No.					
E	WO 00 06737 A (GILBERT CHRISTOPHE FRANCOIS GU; HANSBRO PHILIP MICHAEL (GB); MICRO) 10 February 2000 (2000-02-10) page 2, line 10-12 page 13, line 9 -page 14, line 23 overlap seq.Id 6 page 76, line 31-38 overlap seq.Id 4,10 page 99 -page 100		1-3		
		•			
-			·		
			·		
			•		

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int itional Application No PCT/US 99/30390

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
W0 9818930	A	07-05-1998	AU AU EP EP WO	5194598 A 6909098 A 0942983 A 0941335 A 9818931 A	22-05-1998 22-05-1998 22-09-1999 15-09-1999 07-05-1998
W0 9915675	Α	01-04-1999	AU AU EP	2420800 A 9510598 A 1017828 A	15-06-2000 12-04-1999 12-07-2000
WO 0006737	Α	10-02-2000	NONE		

Form PCT/ISA/210 (patent family annex) (July 1992)

•